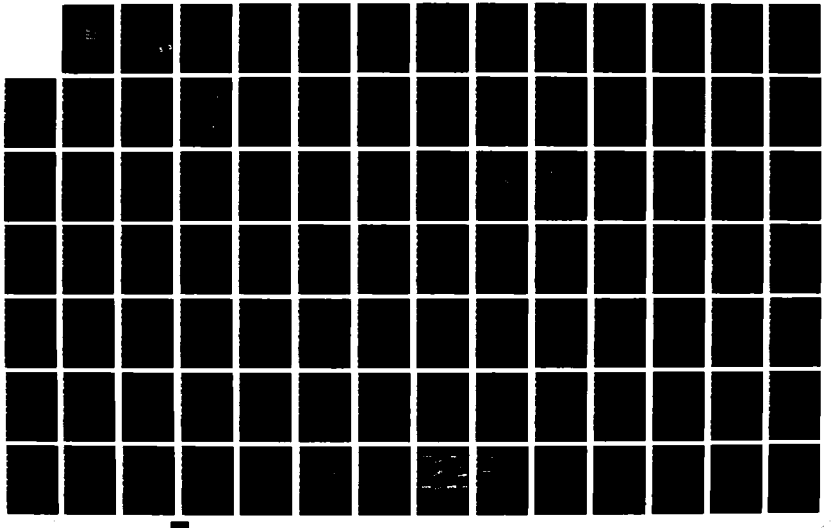
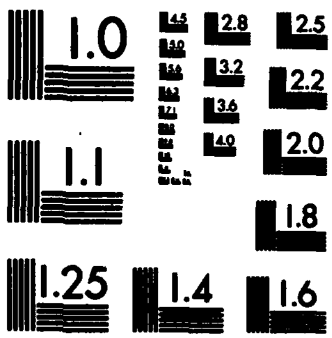
AD-A147 619 ROBUST ADAPTIVE KALMAN TRACKERS FOR SYSTEMS WITH UNKNOWN STEP INPUTS NON-.. (U) WYOMING UNIV LARAMIE R L KIRLIN ET AL. 31 OCT 84 N00014-82-K-0048 1/2

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AD-A147 619

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: None	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Robust Adaptive Kalman Trackers for Systems with Unknown Step Inputs Non-Gaussian Delay Measurement Errors		5. TYPE OF REPORT & PERIOD COVERED: Final Technical Report 2 Nov. 1981-31 Oct. 1984 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): R. Lynn Kirlin and Alireza Moghaddamjoo		8. CONTRACT OR GRANT NUMBER(s): N00014-82-K-0048
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Wyoming, P.O. Box 3295 University Station, Laramie, WY 82071		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 042-470/9-24-81 (410)
11. CONTROLLING OFFICE NAME AND ADDRESS: Chief of Naval Research, Code 410, 800 N. Quincy St. Arlington, VA 22217		12. REPORT DATE: October 20, 1984 13. NUMBER OF PAGES: 111
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): ONR Resident Representative, Rm. 223, Bandelier Hall West, Albuquerque, NM 87131		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release: distribution unlimited.

DTIC ELECTE NOV 19 1984 S D E

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

delay estimation, Kalman estimation, passive sonar, array processing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Signals from target sources produce intersensor delay measurements in sensing arrays. Apriori knowledge of range may improve the delay measurement uncertainty. Correlated inter-sensor noise may be taken into account. Robust statistical methods can be used to further improve the use of the measurements in estimating target states.

84 11 06 05 6

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DTIC FILE COPY

FINAL TECHNICAL REPORT

Project Title: Investigation of effects of model parameters on delay estimates via sequential state estimation.

Contract No.: N00014-82-K-0048

Principal Investigator: Professor R. Lynn Kirlin
Department of Electrical Engineering
University of Wyoming
Laramie, Wyoming 82071
(307) 766-6137

Long Range Scientific Objectives

Passive sonar arrays provide time delay measurements with which the signal sources' location parameters may be estimated. The accuracy of the target location parameters or states is a function of the accuracy of the delay measurements, apriori parameters, number of sensors, intersensor noise correlation, etc. A trade of sensor numbers vs. cost and accuracy as a function of the parameters is desired. Lastly, the effects of non-Gaussian measurement noise on a Kalman tracker are sought and robust, adaptive procedures are considered for improved use of the delays for tracking targets with unknown step acceleration inputs.

Summary

The three year project was divided into three phases. Each of these phases is distinct, but are linked together by the fundamental problem of how to better estimate and use time delays for target location. The four chapters of this report will consist of papers accepted or submitted for publication and which summarize the findings of the three phases.

The first chapter deals with the use of apriori information for correcting delay measurements. The information has to do with wavefront shape, which for distant targets requires a set relation among the time delays. This relation modifies via a Kalman or Bayesian method the delay measurements, and provides a variance decrease.

The second chapter is the result of the second year's effort to understand the effect of intersensor (spatial) noise correlation on the previous results. The Cramer-Rao matrix bound is derived and numerically evaluated for the delay estimates under various conditions of SNR and sensor number, etc.

The third and fourth chapters deal with methods of handling unknown non-Gaussian noise or errors on the delay measurements and unknown process noise and step acceleration inputs to the target state system model. Robust statistical methods are proposed and applied to a first-order linear state system as an example.

List of Technical Reports

Each year's work ended with a technical report. These were submitted to ONR. Their titles are:

1. Analysis of Delay Estimation Improvement Factors Due to Multiple Measurements and A Priori Information, R. Lynn Kirlin and Ernest S. Gale. 30 August, 1982. 46 pp.

2. Optimal Delay Estimation in a Multiple Sensor Array Having Spatially Correlated Noise. R. Lynn Kirlin and Lois A. Dewey. 30 September, 1983. 62 pp.

3. Robust Adaptive Kalman Trackers for Systems with Unknown Step Inputs and Non-Gaussian Delay Measurements Errors. R. Lynn Kirlin and Alireza Moghaddamjoo. 25 October 1984. ~104 pp.

List of Publications

1. R. Lynn Kirlin and Ernest S. Gale, "Analysis of Delay Estimation Improvement Factors due to Multiple Measurements and A Priori Information", accepted for publication in IEEE Trans. ASSP.

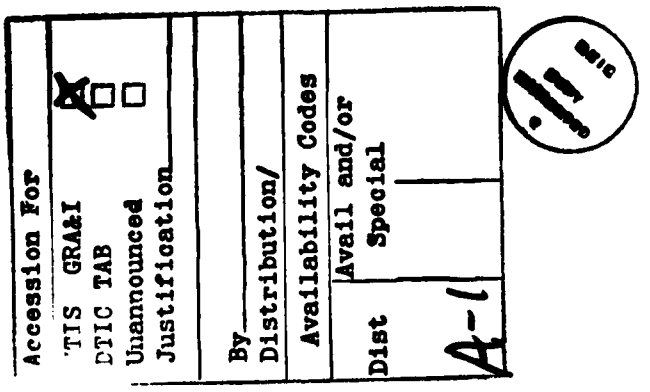
2. R. Lynn Kirlin, "Passive Sonar Delay Estimate Improvement Using A Priori Knowledge and Increased Number of Sensors," in Statistical Signal Processing, E. J. Wegman and J. G. Smith, editors, Marcel Dekker, Inc., New York, 1984, pp. 313-327.

3. R. L. Kirlin, "Optimal Delay Estimation in a Multiple Sensor Array Having Spatially Correlated Noise," IEEE International Conference on Acoustics, Speech, and Signal Processing, 1983, Boston. IEEE Cat. No. 83 CH1841-6, pp. 895-898.

4. R. L. Kirlin and Lois A. Dewey, "Optimal Delay Estimation in a Multiple Sensor Array Having Spatially Correlated Noise," now in second review after revisions, IEEE Trans. ASSP.

5. R. Lynn Kirlin and A. Moghaddamjoo, "Robust Adaptive Kalman Filtering for Systems with Unknown Step Inputs and Non-Gaussian Measurement Errors," submitted to IEEE Trans. ASSP.

6. R. Lynn Kirlin and A. Moghaddamjoo, "A Robust Running-Window Detector and Estimator for Step-Signals in Contaminated Gaussian Noise," submitted to IEEE Trans. ASSP.



CONTENTS

CHAPTER 1

ANALYSIS OF DELAY ESTIMATION IMPROVEMENT FACTORS DUE TO MULTIPLE MEASUREMENTS AND A PRIORI INFORMATION

Abstract

Signals from distant sources produce intersensor delays in towed arrays which may have their measurement variances improved when all possible sensor-pair measurements and **a-priori** knowledge of range are taken into account. Improvement factors are defined and plotted vs. signal-to-noise ratio and vs. number of sensors in the arrays. Clustered vs. equally spaced sensors is another aspect investigated. Analytical results are obtained for three sensors.

I. INTRODUCTION

This paper intends to show the improvement in noisy measurements of passive sonar delays by using a priori range information and an increased number of sensors. The problem is cast into state variable form with the sonar delays as states. Kalman filtering is used to improve, through a priori information and multiple measurements, the delays to be estimated. The extent of improvement can be examined by noting the decrease in the sonar delay estimate variances after one set of measurements is taken and a priori information is taken into account.

The estimation problem is investigated first for three sensors and later expanded to the multiple sensor case. For all cases, the sensors are considered to be perfectly in line, as in an ideal, towed array. When they are clustered, all sensors in a cluster are assumed to share the same spatial point. Other assumptions are that the signal spectrum is the same at every sensor and the noise spectra are also the same, but independent; that is, the noise is independent and identically distributed at each sensor.

With three passive sensors it is possible to use the measured intersensor delays for estimating the bearing and range of a signal source. The measurements are often provided via generalized crosscorrelator [1,6]. The measurements $\underline{z}$ of the delays $\underline{d}$ are generally in error; however, a priori information may improve that error.

Consider the sensor array, delays and measurements depicted in Figure 1.

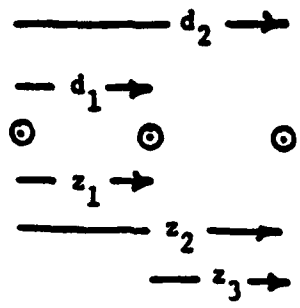


Figure 1. Sensor array, delays d_1 and d_2, measurements z_1, z_2, z_3

The relationship between the measurements and the delays is given by

$$\underline{z} = \underline{H}\,\underline{d} + \underline{e} \tag{1}$$

where

$$\underline{H} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}, \quad \underline{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad \underline{e} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \tag{2}$$

and $\underline{e}$ is the error vector. The crosscorrelator measurement covariance is

$$\mathrm{cov}(\underline{e}) = \underline{R} = r_o^2 \begin{pmatrix} 1 & \gamma & -\gamma \\ \gamma & 1 & \gamma \\ -\gamma & \gamma & 1 \end{pmatrix} \tag{3}$$

where the sign of γ, $0 \le \gamma \le 0.5$, is due to common signals in the cross-correlations of signals from members of sensor pairs. That is, if there is a common signal or sensor in two delay estimates, γ has the positive sign if the common signal is used first in both delays or second in both; but if the common signal is used first in one delay and second in the other, the sign is negative [1]. For large signal-to-noise ratios (SNR) it may be seen from results in [1] that $\gamma \to 1/2$. Similarly for SNR $\to 0$, $\gamma \to 0$.

Using subscripts ij to indicate from i to j, it is clear that $d_{13} = d_{12} + d_{23}$, but how measurements of any of these effect estimation of others has not been clear. In [2] it is shown, for example, that at low SNR, measurement error variance in z_1 can be reduced by a factor of 2/3 by use of z_2 and z_3. Further in [2], a priori knowledge of source (target) range is shown to give an improvement factor of 1/6 at low SNR. The effect of redundant measurements and a priori knowledge is next analyzed for improvement over measurement error covariance in z_2, which improvement was not explored in [2], due to that particular assignment of states. Nor was relative improvement with number of sensors, M, considered.

Now let the a priori uncertainty matrix, cov $\underline{d}_o$, be

$$\underline{P}_o = \sigma_o^2 \begin{pmatrix} 1 & \beta \\ \beta & 4 \end{pmatrix} \tag{4}$$

The parameter β is a weight corresponding roughly to knowledge of target range. For infinite range, $\beta = 2$, because any a priori delay estimation error in $\hat{d}_1(0)$ necessarily is twice as large in $\hat{d}_2(0)$; thus cov $(\hat{d}_1(0), d_2(0)) = 2$ var $\hat{d}_1(0)$. However, when range is totally uncertain, a priori errors are in no way correlated, and β may be set to zero. In any case, $\beta \to 2$ for all reasonable ranges (range much greater than sensor array length). Thus the choice of $\sigma_o^2(4 + \alpha)$ is also reasonable for $p_{22}(0)$.

If σ_o^2 is chosen assuming uniformly distributed bearing, then $\sigma_o^2 = (L/v)^2/2$ where L is the sensor spacing and v is the velocity of sound in the medium, σ_o^2 is the mean square value of $d_1 \simeq (L/v) \sin \theta$, where θ is the counterclockwise angle from broadside.

Measurements of $\underline{d}$ may be improved due to both the dependence among the $\underline{z}$ elements and the *a priori* information through σ_o^2 and β. The correlator measurement variances r_{11} and r_{22} are effectively reduced according to Kalman estimation theory [3] which yields improved or updated estimates of $\underline{d}$,

$$\hat{\underline{d}}_1 = \underline{P}_1 \underline{H}^T \underline{R}^{-1} \underline{Z} \tag{5}$$

where

$$\underline{P}_1 = (\underline{P}_o^{-1} + \underline{H}^T \underline{R}^{-1} \underline{H})^{-1} = \text{cov}(\hat{\underline{d}}_1) \tag{6}$$

The *a priori* $\underline{d}$, $\underline{d}_o$, has been assumed to be the zero vector, and $\underline{d}$ is assumed a constant as for a nonmoving signal source. Extention to more general situations is straight-forward [2, 3]. $\underline{P}_1$ is the new estimate covariance matrix. Thus its element $p_{11}(1)$, for example, is improved (smaller) from $p_{11}(0)$ and also r_{11}. The ratios

$$\eta_1 = \frac{p_{11}(1)}{r_{11}} = p_{11}(1)/r_o^2 \tag{7a}$$

$$\eta_2 = \frac{p_{22}(1)}{r_{22}} = p_{22}(1)/r_o^2 \tag{7b}$$

are the improvement factors which are of interest. They indicate improvement over the correlator measurements.

II. THREE SENSOR ANALYSIS

η_1 and η_2 are functions of σ_o^2, r_o^2, γ, β, and α. This may mean rather complicated expressions, but under certain simplifying conditions they are easily understood. For three sensors, the nondimensional improvement factors are

$$\eta_1 = \frac{\frac{1}{(\sigma_o^2/r_o^2)(4-\beta^2)} + \frac{2}{\gamma+1}}{\frac{1}{(\sigma_o^2/r_o^2)^2(4-\beta^2)} + \frac{2(5-\beta)}{(\sigma_o^2/r_o^2)(4-\beta^2)(\gamma+1)} + \frac{3}{(\gamma+1)^2}} \tag{8}$$

and

$$\eta_2 = \frac{\frac{4}{(\sigma_o^2/r_o^2)(4-\beta^2)} + \frac{2}{\gamma+1}}{\frac{1}{(\sigma_o^2/r_o^2)^2(4-\beta^2)} + \frac{2(5-\beta)}{(\sigma_o^2/r_o^2)(4-\beta^2)(\gamma+1)} + \frac{3}{(\gamma+1)^2}} \tag{9}$$

Next the a priori conditions $\beta = 0$ and $\beta = 2$ correspond to completely uncertain range and infinite range respectively. Letting $\gamma = 0$ or $1/2$ corresponds to SNR = 0 or infinity respectively. The functions η_1 or η_2 are given in Table 1 for these four conditions.

The table shows that for even infinite σ_o^2 or complete a priori uncertainty, improvement is made in $\hat{d}_1$ due to the fact that z_3 contributes more information. That is, for $\sigma_o^2 = \infty$, r_o^2 is reduced by η_1, ranging from 1/6 (infinite range, SNR$\to$0) to 1/4 (infinite range, SNR$\to\infty$) to 2/3 (uncertain range, SNR$\to$0). Only when range is uncertain and SNR$\to\infty$ is there no improvement. These data ($r_o^2/\sigma_o^2 \to 0$) show that both the dependent measurements and a priori knowledge of infinite or large range contribute to improvements.

Analyzing η_2 as $r_o^2/4\sigma_o^2 \to 0$ shows improvements in r_o^2 to $p_{22}(1)$ are only possible for SNR$\to$0, regardless of a priori knowledge of range (β). For both $\beta = 0$ or $\beta = 2$ the improvement is by the factor $\eta_2 = 2/3$. From (5), using (8) and (9), the improved delay estimates are

β	asymptotic γ	η_1	η_2
0 (unc. range)	$\gamma \to 0$ (SNR = 0)	$\dfrac{r_o^2/\sigma_o^2 + 8}{(r_o^2/\sigma_o^2)^2 + 10(r_o^2/\sigma_o^2) + 12}$	$\dfrac{r_o^2/4\sigma_o^2 + 1/2}{(r_o^2/4\sigma_o^2)^2 + (5/2)(r_o^2/4\sigma_o^2) + 3/4}$
2 (inf. range)	$\gamma \to 0$ (SNR = 0)	$\dfrac{1}{r_o^2/\sigma_o^2 + 6}$	$\dfrac{1}{(r_o^2/4\sigma_o^2) + 3/2}$
0 (unc. range)	$\gamma \to 1/2$ (SNR = ∞)	$\dfrac{r_o^2/\sigma_o^2 + 16/3}{(r_o^2/\sigma_o^2)^2 + (20/3)(r_o^2/\sigma_o^2) + 16/3}$	$\dfrac{r_o^2/4\sigma_o^2 + 1/3}{(r_o^2/4\sigma_o^2)^2 + (5/3)(r_o^2/4\sigma_o^2) + 1/3}$
2 (inf. range)	$\gamma \to 1/2$ (SNR = ∞)	$\dfrac{1}{r_o^2/\sigma_o^2 + 4}$	$\dfrac{1}{r_o^2/4\sigma_o^2 + 1}$

Table 1. Measurement variance improvement factors for extreme cases of apriori range (β) and SNR (γ). r_o^2 is measurement variance for d_1; σ_o^2 is apriori variance for d_1 (apriori uncertainty), $4\sigma_o^2$ is apriori variance for d_2.

$$\begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \frac{1}{(\gamma+1)s} \begin{pmatrix} \frac{1}{a} + \frac{2}{\gamma+1} & \frac{\beta}{a} + \frac{1}{\gamma+1} & \frac{\beta-1}{a} - \frac{1}{\gamma+1} \\ \frac{\beta}{a} + \frac{1}{\gamma+1} & \frac{4}{a} + \frac{2}{\gamma+1} & \frac{4-\beta}{a} + \frac{1}{\gamma+1} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \tag{10}$$

where

$$s = \frac{1}{(\sigma_o^2/r_o^2)^2(4-\beta^2)} + \frac{2(5-\beta)}{(\sigma_o^2/r_o^2)(4-\beta^2)(\gamma+1)} + \frac{3}{(\gamma+1)^2} \tag{11a}$$

and

$$a = (\sigma_o^2/r_o^2)(4-\beta^2) \tag{11b}$$

Again under the condition $\beta<2$, and letting $r_o^2/\sigma_o^2 \to 0$ (no a priori knowledge of $\underline{d}$),

$$\begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \quad 0 \leq \beta < 2. \tag{12}$$

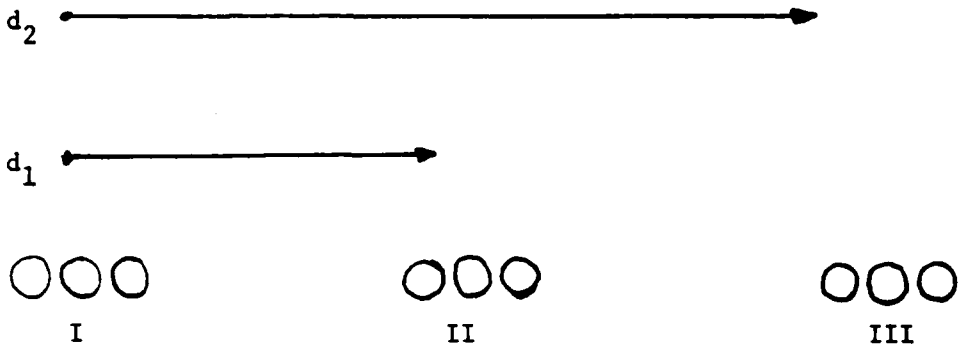
When $\beta = 2$, then the estimates become for $r_o^2/\sigma_o^2 = 0$,

$$\begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \begin{pmatrix} 1/6 & 2/6 & 1/6 \\ 2/6 & 4/6 & 2/6 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \quad \beta = 2 \tag{13}$$

One conclusion is that if all measurement variances are equal, and there is no a priori knowledge of $\underline{d}$, knowledge that range is infinite will improve the error in the measurement of d_2. The improvement is due to interdependence among the measurements and is independent of SNR. Another conclusion is that if anything is known of $\underline{d}$ a priori, σ_o^2 is finite, and measurement errors may be improved with this knowledge also, independent of SNR and a priori range (β) knowledge.

III. MULTI-SENSOR DELAY ESTIMATION GIVEN MEASUREMENT AND A PRIORI COVARIANCES

For optimal range and bearing estimation, the best array configuration [5] for M sensors is three clusters of M/3 sensors each and equal spacing L between groups, as shown in Figure 2.



Equally Spaced Clusters

Figure 2. Multi-sensor Sonar Array Configuration.

All sensors in a "pod" are assumed to be in the same location, i.e., there is no delay between sensors in the same group. The noise spectrum for each sensor is assumed to be the same, but independent from sensor to sensor. With more sensors, more delay measurements are possible, and the increased number of measurements of d_1 and d_2 should improve the Kalman estimate as indicated in Eq. 6. We show, however, that the amount of improvement available depends on the sensors' locations. With three clusters, there are only $(M/3)^2$ measurements instead of $M(M-1)/2$. That is because there are no measurements between sensors in the same cluster. Even though there are fewer measurements, a priori information should produce more improvement, and it does. Also the $(M/3)^2$ measurements are directly end-to-center or end-to-end delays, whereas the $M(M-1)/2$ measurements must be linearly combined to estimate d_1 or d_2.

Many examples from numerical calculation are available in [4], wherein are many contour plots of η for γ vs β. For all numerical results, Q is the ratio r_o^2/σ_o^2.

The η vs. M plots for fixed γ and β give excellent insight as to the improvement of $\hat{d}_1$ and $\hat{d}_2$ by increasing the number of simultaneous delay measurements taken. In general, there is always improvement in the estimate (decrease of η_1) as M increases. In every case, the order of improvement (least to most) is Q=0.1, Q=1, then Q=10. Q=0.1 implies good measurements ($r_0^2 << \sigma_0^2$), and therefore r_0^2 is small. It is more difficult to improve on measurements when they are already very good (Q=0.1) than when they are (relative to a priori) more highly noise contaminated (Q=10).

Figure 3 for three and nine sensors show the amount of improvement available using nine sensors over using just three sensors. Regardless of SNR (indicated through γ), the amount of additional improvement is substantial.

Curves similar to those of Figure 3 were plotted for sensors equally spaced along the line array. These data, Figure 4, show less improvement than that of the clustered sensors as M increases from 3 to 9.

CONCLUSIONS

This paper has examined the effect of using a priori information and multiple measurements on delay estimate variances. The estimate variances after one set of measurements is taken are scaled by the measurement of the delays in question. Following is a quick summary of the major findings in this report and a short description of these results is recorded in Table II.

1. The estimate of d_1 improves greatly due to knowledge that range is much greater than the sensor group spacing; however, $\hat{d}_2$ does not improve over measurement by range information.

2. Improvement factors are plotted for three and nine sensors equally spaced rather than in clusters. Clustering allows even greater improvement from the a priori information.

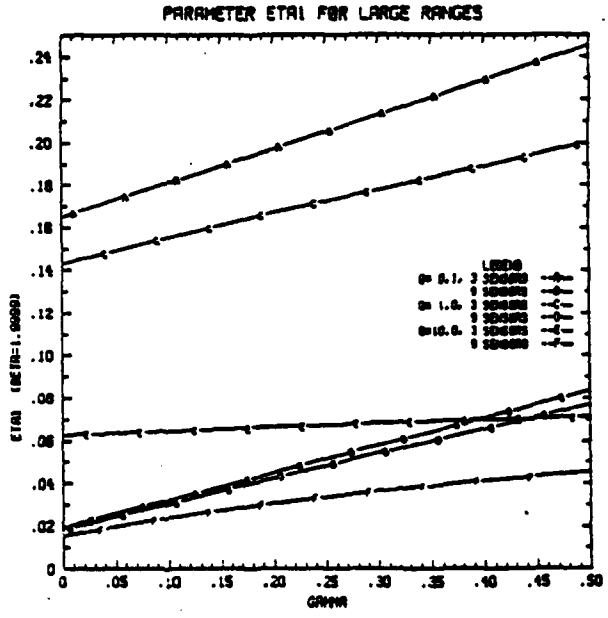


Figure 3. Reduction of η_1 Using Nine Sensors in 3 Clusters Instead of Three Sensors. $Q = r_0^2/\sigma_s^2$.

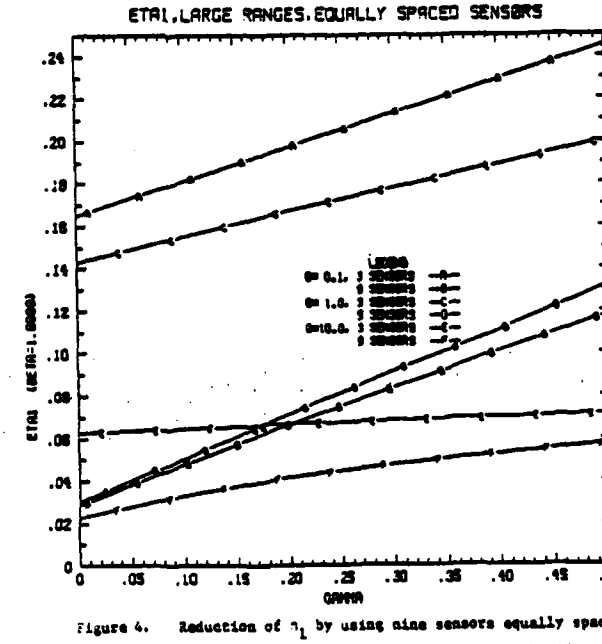


Figure 4. Reduction of η_1 by using nine sensors equally spaced rather than in clusters -- not as much reduction as shown in Figure 3 for clustered

Table II. Improvement Factors η_1 and η_2, Tabulated Qualitative Results.

		SNR → 0 (γ=0)		SNR → ∞ (γ=0.499)		
		apriori range $0 \to \beta \to 2$	# of sensors $3 \to M \to 24$	apriori range $0 \to \beta \to 2$	# of sensors $3 \to M \to 24$	
$Q=0.1$ $\sigma_0^2 >> r_0^2$ good measurements or poor apriori range information		SNR→0 implies very poor measurements. Q=0.1 and SNR→0 are contradictory.		dramatic improvement as $\beta \to 2$	0.0835 at M=9 for $\beta \to 2$	η_1
				no effect	0.330 at M=9 for $\beta \to 2$	η_2
$Q=1.0$ $\sigma_0^2 = r_0^2$ equal measurement and apriori range variances	η_1	dramatic improvement as $\beta \to 2$	0.0401 at M=6 for $\beta \to 2$	steadily increasing improvement as β increases	0.0769 at M=9 for $\beta \to 2$	η_1
	η_2	no effect	0.160 at M=6 for $\beta \to 2$	less improvement as β increases (small M)	0.307 at M=9 for $\beta \to 2$	η_2
$Q=10.0$ $r_0^2 >> \sigma_0^2$ poor measurements or very good apriori range info.	η_1	increasing improvement as β increases (small M)	0.0294 at M=6 for $\beta \to 2$	SNR→∞ implies very good measurements. Q=10.0 and SNR→∞ are contradictory.		
	η_2	less improvement as β increases (small M)	0.117 at M=6 for $\beta \to 2$			

REFERENCES

1. W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation". Journal of the Acoustic Society of America, vol. 58, no. 1, July 1975, pp. 201-207.

2. R. L. Kirlin, D. F. Moore, and R. F. Kubichek, "Improvement of delay measurements from sonar arrays via sequential state estimation", IEEE Trans. ASSP, vol. ASSP-29, no. 3, June 1983, Part II, pp. 514-518.

3. A. Gelb, edition, Applied Optimal Estimation, M.I.T. Press, Cambridge, Mass., 1974.

4. E. S. Gale, "Passive sonar delay estimate improvement using a priori knowledge and increased number of sensors, "M.S.E.E. Thesis, University of Wyoming, July 1982.

5. G. C. Carter, "Variance bounds for passively locating an acoustic source with symmetric line array", Journal of the Acoustical Society of America, Vol. 62, No. 4; Oct. 1977, pp. 922-926.

6. C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay, "IEEE Trans. Acoust., Speech, and Signal Proc., Vol. ASSP-24, No. 4, pp. 320-327, 1976.

CHAPTER 2

OPTIMAL DELAY ESTIMATION IN A MULTIPLE SENSOR ARRAY HAVING SPATIALLY CORRELATED NOISE

Abstract

The maximal likelihood (ML) estimation of time-of-arrival differences for signals from a single source or target arriving at $M \geq 2$ sensors has been the subject of a large number of papers in recent years. These time differences or delays enable target location. Nearly all previous work has assumed noises which are independent among all sensors. Herein noises are taken to have complex correlation between sensors. A set of nonlinear equations in the unknown delays is derived. The Fisher information matrix (FIM) for the estimates is also derived. The Cramer Rao Matrix Bound (CRMB), which is the inverse of FIM, shows optimal estimator covariances. Computer evaluations are given for CRMB elements with varied SNR and noise covariance values typical of turbulent boundary layer noise in towed arrays.

OPTIMAL DELAY ESTIMATION IN A MULTIPLE SENSOR ARRAY HAVING SPATIALLY CORRELATED NOISE[+]

I. Introduction

The estimation of time-of-arrival differences for signals from a single source or target arriving at multiple sensors has been the subject of a considerable number of papers in recent years. These time delay differences, or simply delays, enable target localization through straightforward geometrical considerations when the signal path is non-dispersive [1,2]. Although target location is the primary goal, delay estimation is essentially equivalent as there is a one-one*, although nonlinear, relation between the maximum-likelihood (ML) delay vector and ML location vector.

Essentially all of the results of available literature (except [9]) have been based not only upon the geometric and non-dispersive assumptions stated above but also upon noise spectra which are independent among sensors. The independent noise assumption is adequate if either the sensor self-noise is dominant or the sensors are spatially separated sufficiently such that the environmental noise is indeed independent or uncorrelated among sensors. However, this is not always a reasonable assumption and the effects of spatially correlated noise in the estimation of delays and delay variances must be considered. Thus appropriate analyses are herein undertaken to consider correlated noise. Results are compared to those previously published for independent sensor noises.

Owsley and Fay [11] have considered correlated noise when clustering sensors and optimizing beamformers. The comparable optimization of delay estimation has not previously been approached. By choosing the correlation parameter ρ, we may include the proportionality of correlated turbulent boundary layer tow-noise and isotropic sensor noise.

The basic approach is to assume that complex Fourier coefficients $X_i(k)$ at the $i^{\underline{th}}$ sensor for the $k^{\underline{th}}$ frequency are available, having been obtained from T-second time records, where T is long with respect to the signal correlation time.

---

* For an array with three sensors in line there is an ambiguity in the sign of bearing angle, which we assume may be solved with additional information.

+ This study funded under office of Naval Research, contract number N00014-82-K-0048.

The time data records are

$$x_i(t) = s(t-d_i) + n_i(t), \quad i = 1,2,\ldots M. \tag{1}$$

where d_i are the delays from the reference sensor to the $i^{\underline{th}}$ sensor ($d_1 = 0$), $s(t)$ is a zero-mean, Gaussian, stationary signal, and $n_i(t)$ is the additive Gaussian noise at the $i^{\underline{th}}$ sensor.

II. Background

The problem of delay vector estimation for multiple sensors has been studied with the above approach in original papers by Hahn and Tretter [3], Hahn [4] and Schultheiss [5]. Closely following their presentations, let

$$X_i(k) = \frac{1}{T} \int_{-T/2}^{T/2} x_i(t) \exp\{-jk\omega_o t\}dt, \quad k = 1,2,\ldots,K, \tag{2}$$

where $\omega_o = 2\pi/T$. Define a vector X containing the above MK Fourier coefficients as elements. If $S(\omega)$ and $N_i(\omega)$ are the signal and noise spectra at the $i^{\underline{th}}$ sensor, the probability density for x can be written

$$p(\underline{X}) = (\pi^{MK} \prod_{k=1}^{K} \det R(k))^{-1} \exp[-\sum_{k=1}^{K} X^T(k)R^{-1}(k)X^*(k)] \tag{3}$$

where

$$X(k) = [X_1(k), X_2(k),---,X_M(k)]^T$$

$$\underline{X} = [X^T(1); X^T(2),---,X^T(K)]^T$$

$$V(k) = [1, \exp(-jk\omega_o d_2),---,\exp(-jk\omega_o d_M)]^T$$

$N(k) = [N_{ij}(k)]$, an MxM matrix of noise cross-power spectra

$$R(k) = N(k) + S(k)\, V^*(k)V^T(k)$$

and where * superscript denotes complex conjugation.

In order to obtain the ML estimate of delays, determinant and inverse theorems of use are

$$|R| = |N + SV^*V^T| = |N||I + N^{-1}SV^*V^T| = |N|(1 + SV^TN^{-1}V^*) \tag{4}$$

and

$$R^{-1} = N^{-1} - N^{-1} V^*(V^T N^{-1} V^* + 1/S)^{-1} V^T N^{-1} \tag{5}$$

Defining elements of N^{-1} as N^{ik}, the likelihood function of the delay vector $D^T = (d_2, d_3, \ldots, d_M)$ is, using (4),

$$\Lambda = \ell n\, p(\underline{X}) = -\ell n(\pi^{MK}) - \sum_{B+} \ell n[|N|\ (1 + SV^T N^{-1} V^*)]$$

$$- \sum_{B+} [X^T N^{-1} X^* - \frac{X^T N^{-1} V^* V^T N^{-1} X^*}{V^T N^{-1} V^* + 1/S}] \tag{6}$$

where $\sum_{B+}$ means sum over positive frequencies.

Hahn and Tretter [3] have shown that, when N is diagonal, $[N_1, N_2 \ldots, N_M]$, the Fisher information matrix for D (FIM = - <grad<grad $\ell n\, \Lambda$>> where <•> is expected value) is

$$FIM = \sum_{B+} 2\omega^2 \frac{S^2}{1 + \sum S/N_i} [(\text{tr } N^{-1}) N_p^{-1} - N_p^{-1} 1 1^T N_p^{-1}] \tag{7}$$

where N_p^{-1} is N^{-1} with the first row and column removed. The Cramer-Rao Matrix bound (CRMB) for the delays D is $(FIM)^{-1}$. The ML estimate covariance is known to asymptotically approach the CRMB. The ML estimate for small delays ($\hat{D}$ is the error when D = 0) and independent noise is

$$\hat{D} = -<C>^{-1} B^T, \tag{8}$$

where

$$<C>^{-1} = FIM, \tag{9}$$

$$B = \sum_{B+} j\omega \frac{S}{1 + \sum S/N_i} 1^T N^{-1} [X X_p^{*T} - X^* X_p^T] N_p^{-1}, \tag{10}$$

and x_p is X(k) with the first element ($X_1(k)$) removed.

Hahn and Tretter also show that the ML D estimate can be implemented either as a beamformer (ideally in real time only when the N_i are proportional, because of phase-matching filter criteria), or as a cross correlator

system which produces the M(M-1)/2 delay estimates. The correlator system has cross-spectral filters

$$|F_{ij}|^2 = \frac{S/N_iN_j}{1 + \sum_k S/N_k} \quad , \; i, j = 1, 2, \ldots, M, \quad i \neq j. \tag{11}$$

The error covariance matrix for the pair-wise delay estimates of the correlators is shown by Hahn [4] to have elements

$$\text{var } (\hat{d}_{ij}) = \frac{2\pi}{T} \frac{\int_B \omega^2 |F_{ij}|^4 [N_iN_j + S(N_i + N_j)] d\omega}{(\int_B \omega^2 |F_{ij}|^2 \, Sd\omega)^2} \tag{12}$$

$$\text{covar}(\hat{d}_{ij}, \hat{d}_{k\ell}) = 0, \; i, j, k, \ell \text{ all distinct}$$

$$\text{covar}(\hat{d}_{ij}, \hat{d}_{i\ell}) = \frac{2\pi}{T} \frac{\int_B \omega^2 |F_{ij}|^2 |F_{i\ell}|^2 \, SN_i d\omega}{\int_B \omega^2 |F_{ij}|^2 Sd\omega \int_B \omega^2 |F_{i\ell}| Sd\omega} \quad , \; i \neq \ell \tag{13}$$

$$= - \text{covar } (\hat{d}_{ij}, \hat{d}_{\ell i}) \quad , \; j \neq \ell$$

It is emphasized that these are correlator error covariances of the $\hat{d}_{ij}$ and <u>not</u> ML estimator error variances, which are derived herein.

The delays having covariance matrix defined by (12) and (13) are not the M-1 delays referred to a single sensor. Hahn and Tretter have shown how to use weighted linear combinations of the M(M-1)/2 cross correlation delay estimates, $\hat{d}_{ij}$, to form an estimate for $D = (d_i)$ which achieves the CRMB of (7).

With independent noises maximization of Λ in (6) over the vector D concentrates on the second term in the second summation, because other terms are not dependent on the d_i. This is not generally the case, and an analytical solution is not available, as was pointed out in the multipath analysis given by Owsley [6]. However, the generation of a set of nonlinear equations in the unknowns d_{ij} may be obtained.

In the next section ML estimator equations for the M(M-1)/2 pairwise delays are derived. Section IV produces the CRMB for a subset of these having M-1 independent delays. Section V considers the M-1 delays $d_i - d_1$, and Section VI derives their CRMB.

III. Estimation of an Independent Subset of the M(M-1)/2 Delays

This section will determine equations for ML estimates of an independent subset of the M(M-1)/2 delays $d_i - d_k = d_{ik}$. In contrast most other papers referenced find ML estimates of either the M-1 delays $(d_i - d_1)$, $2 \le i \le M$, or other parameters such as range and bearing, functions of which the d_i may be written. The reason for our choosing the d_{ik} is that a-priori information about linear relationships among them may subsequently be used as in [10] to improve the delay estimates $\hat{d}_i - d_1$ or any other subset.

Consider the two summations in (6), the only functions of D. We would like to solve for the $\hat{d}_{ik}$ which maximizes

$$\Lambda_1 = \sum_{B+} \left[-\ell n(g) + \frac{X^T N^{-1} V^* V^T N^{-1} X^*}{g}\right] \tag{14}$$

$$g = 1/S + V^T N^{-1} V^* = 1/S + \sum_p \sum_{q>p} (\cos \omega d_{pq} \, \mathrm{Re}\{N^{pq}\} + \sin \omega d_{pq} \, I_m\{N^{pq}\}) \tag{15}$$

Differentiating Λ_1 with respect to d_{ik} (assuming all d_{ik} independent) and setting this equal to zero and rearranging gives

$$\sum_{B+} \left\{\frac{\partial (V^T N^{-1} V^*}{\partial d_{ik}} \frac{1}{g} + X^T N^{-1} \, U \, N^{-1} X^*\right\} = 0 \tag{16}$$

The matrix $U(i,k) = (u_{mn}(i,k))$ has typical elements with values which differ according to whether or not $(m,n) = (i,k)$ or (k,i). These are found to be

$$u_{mn} = \frac{1}{g} \begin{cases} e^{j\omega d_{ik}} (j\omega - g_1/g) & ,(m,n)=(i,k) \\ e^{-j\omega d_{ik}} (-j\omega - g_1/g) & ,(m,n)=(k,i) \\ -e^{j\omega d_{mn}} \, g_1/g & ,(m,n)\neq(i,k),\ (k,i) \end{cases} \tag{17}$$

Insertion of (17) into (16) constitutes (M-1) equations, nonlinear, to be solved for the d_{ik}. Note that only M-1 delays can be independent. We now turn our attention to the CRMB.

IV. The Cramer-Rao Bound for an Independent Subset of the M(M-1)/2 Delay Estimates of d_{ik}

As is well known, maximum likelihood estimators have variances which approach the Cramer-Rao bound. The variance bounds for the $\hat{d}_{ik}$ are the elements in the diagonal of

$$\text{CRMB} = (\text{FIM})^{-1} = (-\langle \text{grad}(\text{grad } \ell n \ \Lambda_1)^T \rangle)^{-1}, \tag{18}$$

wherein FIM is the Fisher Information Matrix, grad Λ_1 is a row vector whose m^{th} element is the derivative of Λ_1 with respect to the m^{th} delay (the m^{th} d_{ik} here), Λ_1 is the expression in Eq. (14), and $\langle \cdot \rangle$ denotes expectation. The outer gradient operator creates a matrix whose elements are $\frac{\partial}{\partial d_{rt}} (\frac{\partial \Lambda_1}{\partial d_{ik}})$. We have already found the inner partial -- the result is Eq. (16). For any M-1 independent delays the following applies.

Taking and negating the second partial with respect to d_{rt} gives

$$\sum_{B+} \frac{1}{g} (g_{12} - g_1 g_2/g - \frac{1}{g} \sum_{B+} X^T N^{-1} B N^{-1} X^* \tag{19}$$

where

$$g_1 = \frac{\partial g}{\partial d_{ik}} = 2\omega(-\text{Re}\{N^{ik}\}\sin\omega d_{ik} + \text{Im}\{N^{ik}\}\cos\omega d_{ik}) \tag{20}$$

$$g_2 = \frac{\partial g}{\partial d_{rt}} = 2\omega(-\text{Re}\{N^{rt}\}\sin\omega d_{rt} + \text{Im}\{N^{rt}\}\cos\omega d_{rt}) \tag{21}$$

and where b_{mn} has the following values (letting $\mu = 0$ if $(r,t) \neq (i,k)$, $\mu = 1$ if $(r,t) = (i,k)$)

<u>$(m,n) = (i,k), (k,i)$</u>

$$b_{ik} = (\frac{-g_2}{g} j\omega - \mu(\omega^2 + \frac{g_{12}}{g} + \frac{g_1}{g} j\omega) + 2\frac{g_1 g_2}{g^2}) e^{j\omega d_{ik}} \tag{22a}$$

$$b_{ki} = b_{ik}^*$$

<u>$(m,n) = (r,t), (t,r)$</u> ; $(r,t) \neq (i,k)$

$$b_{rt} = \frac{-g_1}{g} j\omega + \frac{2g_1 g_2}{g^2} \tag{22b}$$

$$b_{tr} = b_{rt}^*$$

$\underline{(m,n) \neq (i,k),\ (k,i),\ (r,t),\ (t,r)}$

$$b_{mn} = \left(2\,\frac{g_1 g_2}{g^2} - \mu\,\frac{g_{12}}{g}\right) e^{j\omega d_{mn}} \tag{22c}$$

Using

$$\langle X_r X_t^* \rangle = \begin{cases} S e^{-j\omega d_{rt}} + N^*_{rt}, & r \neq t \\ S + N_1, & r = t \end{cases} \tag{23}$$

the elements of the FIM are

$$\begin{aligned}
(\mathrm{FIM})_{rt,ik} = {} & \sum_{B+} \frac{1}{g}\,(g_{12} - g_1 g_2/g) \\
& - \sum_{B+} \frac{1}{g} \Big\{ \sum_{m} b_{mm} \big[(S+N_1)(N^{mm})^2 \\
& + 2 \sum_{q \neq m} \big((S\cos\omega d_{mq} + \mathrm{Re}\{N^*_{mq}\})\,\mathrm{Re}\{N^{mm} N^{mq}\} \\
& \quad - (-S\sin\omega d_{mq} + \mathrm{Im}\{N^*_{mq}\})\,\mathrm{Im}\{N^{mm} N^{mq}\} \big) \\
& + 2 \sum_{p \neq m} \sum_{q > p} \big((S\cos\omega d_{pq} + \mathrm{Re}\{N^*_{pq}\})\,\mathrm{Re}\{N^{pm} N^{mq}\} \\
& \quad - (-S\sin\omega d_{pq} + \mathrm{Im}\{N^*_{pq}\})\,\mathrm{Im}\{N^{pm} N^{mq}\} \big) \\
& + \sum_{p \neq m} |N^{pm}|^2 \,(S\cos\omega d_{pm} + \mathrm{Re}\{N^*_{pm}\}) \big] \\
& + 2 \sum_{m} \sum_{n > m} \big[\mathrm{Re}\{b_{mn}\} \langle \mathrm{Re}\{G(m,n)\} \rangle \\
& \quad - \mathrm{Im}\{b_{mn}\} \langle \mathrm{Im}\{G(m,n)\} \rangle \big] \Big\}
\end{aligned} \tag{24}$$

where

$$\langle \mathrm{Re}\{G(m,n)\} \rangle = \sum_{p} \sum_{q} \big(\mathrm{Re}\{N^{pm} N^{nq}\}\,(S\cos\omega d_{pq} + \mathrm{Re}\{N^*_{pq}\}) - \mathrm{Im}\{N^{pm} N^{nq}\}\,(-S\sin\omega d_{pq} + \mathrm{Im}\{N^*_{pq}\}) \big) \tag{25a}$$

and

$$\langle \mathrm{Im}\{G(m,n)\} \rangle = \sum_{p} \sum_{q} \big(\mathrm{Re}\{N^{pm} N^{nq}\}\,(-S\sin\omega d_{pq} + \mathrm{Im}\{N^*_{pq}\}) + \mathrm{Im}\{N^{pm} N^{nq}\}\,(S\cos\omega d_{pq} + \mathrm{Re}\{N^*_{pq}\}) \big) \tag{25b}$$

Use of these elements in the FIM is restricted for inversion to the CRMB to M-1 independent delays.

We now investigate for the M-1 d_i changes in earlier results caused by consideration of noise which is spatially correlated.

V. Maximum Likelihood Estimation of the M-1 Delays $d_i - d_1$

We again maximize Λ by maximizing

$$\Lambda_1 = -\sum_{B+} \ln g + \sum_{B+} X^T N^{-1} V^* V^T N^{-1} X^*/g \tag{26}$$

Writing

$$X^T N^{-1} = (\sum X_p N^{p1}, \sum X_p N^{p2}, \ldots, \sum X_p N^{pm}) \tag{27}$$

and the $m,n^{\underline{th}}$ element

$$(V^* V^T)_{m,n} = v_{mn} = e^{j\omega(d_m - d_n)}$$

gives

$$\begin{aligned} X^T N^{-1} V^* V^T N^{-1} X^*/g &= \frac{1}{g} \sum_m \sum_n v_{mn} (X_m X_n^* N^{mm} N^{nn} + X_m N^{mm} \sum_{q \neq n} X_q^* N^{qn} \\ &\quad + X_n^* N^{nn} \sum_{p \neq m} X_p N^{pm} \\ &\quad + \sum_{p \neq m} \sum_{q \neq n} X_p X_q^* N^{pm} N^{qn}). \end{aligned} \tag{28}$$

In this form it may be seen that (26) differs from the spatially uncorrelated noise case only in the $-\sum_{B+} \ln g$ term and the terms in parentheses in (28) other than $X_m X_n^* N^{mm} N^{nn}$. If $\rho = 0$, $\ln g$ is not a function of the delays and all N^{pq}, $p \neq q$, are zero. Then as the literature cited shows [3,4], maximization of (6) reduces to either a beamformer (choosing M-1 d_i) or a system of $M(M-1)/2$ correlators (choosing $d_i - d_k$).

Maximizing Λ means solving for d_i in $\partial\Lambda_1/\partial d_i = 0$.

Using

$$g_i = \frac{\partial g}{\partial d_i} = 2\omega \sum_{p \neq i} (-\mathrm{Re}\{N^{ip}\}\sin\omega(d_i - d_p) + \mathrm{Im}\{N^{ip}\}\cos\omega(d_i - d_p)) \quad (29)$$

gives

$$\frac{\partial \Lambda}{\partial d_i} = -\sum_{B+} \frac{2\omega}{g} \sum_{p \neq i} (\mathrm{Re}\{N^{ip}\}\sin\omega(d_i - d_p) + \mathrm{Im}\{N^{ip}\}\cos\omega(d_i - d_p)) + \sum_{B+} X^T N^{-1} A N^{-1} X^* \quad (30)$$

where $A = (a_{mn}(i))$ and

$$a_{mn}(i) = \frac{1}{g} \begin{cases} -g_i e^{j\omega(d_m - d_n)}/g); & m,n \neq i \text{ or } m = n = i \\ (j\omega - g_i/g)\, e^{j\omega(d_i - d_n)}; & m = i,\ n \neq i \\ (-j\omega - g_i/g)\, e^{-j\omega(d_i - d_m)}; & m \neq i,\ n = i \end{cases} \quad (31)$$

Because the $a_{mn}(i)$ are functions of g_i and g, and g_i and g are functions of all delay differences $d_p - d_q$, the solution for d_i cannot be found in terms of X_i and X_1 alone nor even as a linear combination of the $X_p X_q^* e^{j\omega(d_p - d_q)}$ correlators.

VI. The Cramer-Rao Matrix Bound for the M-1 Delays $d_i - d_1$.

It is well known that ML estimators approach the Cramer-Rao bound (CRMB). The variance bounds for the delay estimates $\hat{d}_i$ are the diagonal elements in

$$\mathrm{CRMB} = (\mathrm{FIM})^{-1} = (-\langle \mathrm{grad}\ (\mathrm{grad}\ \Lambda_1)^T \rangle)^{-1}, \quad (32)$$

wherein FIM is the Fisher Information Matrix, grad Λ_1 is a row vector whose m^{th} element is the derivative of Λ_1 with respect to the m^{th} delay d_{m+1}, Λ_1 is the expression in (14), and $\langle\cdot\rangle$ denotes expectation. The outer gradient operator creates a matrix whose elements are $-\frac{\partial}{\partial d_k}(\frac{\partial \Lambda}{\partial d_i})$. The inner partial has already been given by (30) and (31). Continuing with $\frac{\partial}{\partial d_k} = g_k$ and

$$g_{ik} = \frac{\partial g_i}{\partial} = 2\omega^2(\mathrm{Re}\{N^{ik}\}\cos\omega(d_i - d_k) + \mathrm{Im}\{N^{ik}\}\sin\omega(d_i - d_k)) \quad (33)$$

We find

$$\frac{-\partial}{\partial d_k}\left(\frac{\partial \Lambda_1}{\partial d_i}\right) = \sum_{B+} \frac{1}{g}\left[g_{ik} - g_i g_k/g\right]$$

$$-\sum_{B+} \frac{1}{g} X^T N^{-1} U N^{-1} X^* \tag{34}$$

where u_{mn} have the following values:

For $i \neq k$, $m \neq i,k$; $n \neq i,k$

$$u_{mn} = \left(\frac{-g_{ik}}{g} + \frac{2g_i g_k}{g^2}\right) e^{j\omega(d_m - d_n)} \tag{35a}$$

$(m,n) = (k,i)$ or (i,k)

$$u_{ki} = \left(\omega^2 - \frac{g_{ik}}{g} + \frac{2g_i g_k}{g^2} + j\omega\left(\frac{g_k}{g} - \frac{g_i}{g}\right)\right) e^{j\omega(d_k - d_i)} \tag{35b}$$

$$u_{ik} = u_{ki}^*$$

$m = k$; $n \neq i,k$ or $n = k$; $m \neq i,k$

$$u_{kn} = \left(-\frac{g_{ik}}{g} + \frac{2g_i g_k}{g^2} - \frac{g_i}{g} j\omega\right) e^{j\omega(d_k - d_n)} \tag{35c}$$

$$u_{mk} = u_{km}^*$$

$m = i$; $n \neq i, k$ or $n = i$; $m \neq i,k$

$$u_{in} = \left(\frac{-g_{ik}}{g} + \frac{2g_i g_k}{g^2} - \frac{g_k}{g} j\omega\right) e^{j\omega(d_i - d_n)} \tag{35d}$$

$$u_{mi} = u_{im}^*$$

$m = n$

$$u_{mm} = \left(\frac{-g_{ik}}{g} + \frac{2g_i g_k}{g^2}\right) \tag{35e}$$

giving

$$(FIM)_{ik} = \sum_{B+} \frac{1}{g} (g_{ik} - g_i g_k/g)$$

$$- \frac{1}{g} \sum_{B+} \{ \sum_{m} u_{mm} [(S+N_1)(N^{mm})^2$$

$$+ 2 \sum_{q \neq m} (\mathrm{Re}\{N^{mm} N^{mq}\} (S \cos\omega(d_m - d_q) + \mathrm{Re}\{N^*_{mq}\}$$

$$- \mathrm{Im}\{N^{mm} N^{mq}\} (-S \sin\omega(d_m - d_q) + \mathrm{Im}\{N^*_{mq}\}))$$

$$+ 2 \sum_{p \neq m} \sum_{q > p, \neq m} (\mathrm{Re}\{N^{pm} N^{mq}\} (S \cos\omega(d_p - d_q) + \mathrm{Re}\{N^*_{pq}\})$$

$$- \mathrm{Im}\{N^{pm} N^{mq}\} (-S \sin\omega(d_m - d_q) + \mathrm{Im}\{N^*_{mq}\}))$$

$$+ \sum_{p \neq m} |N^{pm}|^2 (S + N_1)]$$

$$+ 2 \sum_{m} \sum_{n > m} [\mathrm{Re}\{u_{mn}\} < \mathrm{Re}\{G(m,n)\} >$$

$$- \mathrm{Im}\{u_{mn}\} < \mathrm{Im}\{G(m,n)\} >] \} \qquad (36)$$

$$\frac{-\partial^2 \Lambda_1}{\partial d_i^2} = \sum_{B+} \frac{1}{g} \{[g_{ii} - g_i^2/g] - X^T N^{-1} W N^{-1} X^*\} \qquad (37)$$

where w_{mn} have the following values.

<u>$m \neq i$, $n \neq i$, $m \neq n$</u>

$$w_{mn} = (-g_{ii}/g)\, e^{j\omega(d_m - d_n)} + 2(g_i^2/g^2)\, e^{j\omega(d_m - d_n)} \qquad (38a)$$

<u>$m = i$, $n \neq i$ or $n = i$, $m \neq i$</u>

$$w_{in} = (-\omega^2 - g_{ii}/g + 2\, g_i^2/g^2 - 2(g_i/g)j\omega)\, e^{j\omega(d_i - d_n)} \qquad (38b)$$

$$w_{mn} = w_{im}^*$$

<u>$m = n$</u>

$$w_{mm} = -\, g_{ii}/g + 2\, g_i^2/g^2 \qquad (38c)$$

Using the above results gives

$$(FIM)_{ii} = \sum_{B+} \frac{1}{g} (g_{ii} - g_i^2/g)$$

$$
\begin{aligned}
&-\frac{1}{g}\sum_{B+}\Big\{\sum_{m} w_{mm}\,[(S+N_1)(N^{mm})^2 \\
&\quad + 2N^{mm}\sum_{q\neq m}\big((S\cos\omega(d_m-d_q) + \mathrm{Re}\{N^*_{mq}\})\mathrm{Re}\{N^{mq}\} \\
&\qquad -(-S\sin\omega(d_m-d_q) + \mathrm{Im}\{N^*_{mq}\})\,\mathrm{Im}\{N^{mq}\}\big) \\
&\quad + 2\sum_{p\neq m}\sum_{q>p,\neq m}(S\cos\omega(d_p-d_q) + \mathrm{Re}\{N^*_{pq}\})\,\mathrm{Re}\{N^{pm}N^{mq}\} \\
&\qquad -(-S\sin\omega(d_p-d_q) + \mathrm{Im}\{N^*_{pq}\})\,\mathrm{Im}\{N^{pm}N^{mq}\}\big) \\
&\quad + \sum_{p\neq m}|N^{pm}|^2\,(S+N_1)] \\
&\quad + 2\sum_{m}\sum_{n>m}[\mathrm{Re}\{w_{mn}\}<\mathrm{Re}\{G(m,n)\}> \\
&\qquad -\mathrm{Im}\{w_{mn}\}<\mathrm{Im}\{G(m,n)\}>]\Big\}
\end{aligned}
\tag{39}
$$

To compare with previous results observe in (36) and (37) that if noise is spatially uncorrelated, $g_i = 0$, and only $u_{ik} = u^*_{ki} = \omega^2 e^{\pm j\omega(d_i-d_k)}$ and $w_{in} = w^*_{ni} = -\omega^2 e^{\pm j\omega(d_i-d_n)}$ are non-zero. Further, in (25) $p = m$ and $q = n$ give the only non-zero terms. Utilizing the above,

$$
\begin{aligned}
(\mathrm{FIM})_{ik} &= -2\sum_{B+}\frac{\omega^2}{g}\,(\cos\omega(d_i-d_k)\,N^{ii}N^{kk}\,S\cos\omega(d_i-d_k) \\
&\qquad + \sin\omega(d_i-d_k)\,N^{ii}N^{kk}\,S\sin\omega(d_i-d_k)) \\
&= -2\sum_{B+}\frac{\omega^2}{g}\,N^{ii}N^{kk}S,
\end{aligned}
$$

where $g = \frac{1}{S} + \sum_i N^{ii}$,

and similarly

$$
(\mathrm{FIM})_{ii} = 2\sum_{B+}\sum_{n\neq i}\frac{\omega^2}{g}\,N^{ii}N^{nn}S.
$$

It is readily seen that this FIM is identical to Eq. 7 (the same as Eq. 12 in [3]).

Unfortunately the FIM defined by (36) and (39) has elements which are in general functions of the delays themselves, making analysis difficult. However, in the next section we will assume a signal source at infinity, allowing some simplification.

VII. Evaluation of the CRMB

It is unreasonable to evaluate and invert the FIM in sections IV or VI in general because it is a function of all d_i. However, if wavefront curvature is ignored, each delay may be written $d_i = i\Delta$ where Δ is the delay between adjacent sensors. We may also let Δ vary between zero and $\omega\Delta = \pi$ for a single frequency. Then $d_p - d_q = (p-q)\Delta$ for example. This is the beamformer case.

Because of the generality of the formulas we may also vary the elements of N, using the symmetric matrix (as in [11])

$$N = N_1 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{m-1} \\ & 1 & \rho_1 & \cdots & \rho_{m-2} \\ & & 1 & \cdots & \vdots \\ & & & & 1 \end{pmatrix}$$

wherein $\rho_r = \rho_o e^{-r\sigma} e^{j\omega\beta}$. Such a correlation is appropriate for turbulent boundary layer noises and its magnitude with respect to the unity diagonal accounts for isotropic noise. In the following simulations we choose $|\rho_1| =$ 0, 0.2, 0.4 and $\omega\beta$ having values 0 through $\pi/2$.

Figures 2-10 show the CRMB (1,1) element, center element, or last element as a function of the various parameters. Table I is presented as a guide to comparisons.

The formulas for the FIM may be applied to arrays with clustered elements as well, if the spacing between clusters is considered. We have done this in producing the data in Figures 11 and 12. Zero correlation between clusters is assumed.

The clustered (or grouped) arrays studied are as shown in Figure 1.

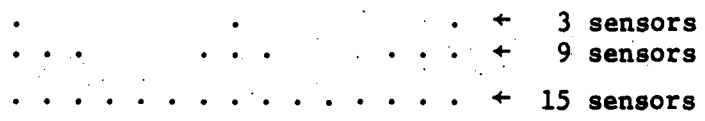


Figure 1. Scheme for clustering sensors

The spacing between array ends and ends-to-center remains fixed. The effect of adding sensors to the cluster when spatially correlated noise is present can then be observed.

Table I. Guide to Figure Variables

Fig #	Sensors M	Measure-ments	SNR	ρ	ωΔ	Ordinate CRMB element	Abscissa	Notes
2	3	2		6 values	0	1,1	SNR	
3	3	2		"	1.571	1,1	SNR	
4	3	2	0.1	0.0,0.2,0.4		1,1	ωΔ	
5	3	2	0.1	"		2,2	ωΔ	
6	3	2	10.0	"		1,1	ωΔ	
7	15	14	0.1	"		1,1	ωΔ	
8	15	14	0.1	"		7,7	ωΔ	
9	15	14	0.1	"		14,14	ωΔ	
10			1.0	"	0	1,1	sensors,M	equi-spaced array
11			0.1,1.0,10	0.2	0	M-1,M-1	sensors,M	grouped array
12			1.0	0.0,0.2,0.4	0	M-1,M-1	sensors,M	"

Comments derived from the Figures are as follows

1. Figures 2 and 5 show that variance decreases monotonically with SNR and that variations in ρ from 0 through 0.4 /.45° have unpredictable, but not large effects.
2. Figures 4 and 7 for example show that more sensors (from 3 to 15) will reduce the variance of a delay.
3. Figures 8 and 9 show that variance bounds for delays end-to-center will vary with ρ differently than those for end-to-end, but not a lot. Also the end-to-end delays vary least.
4. Comparing Figs. 1 and 3 for example shows that the effect of ρ on a delay estimator will vary with $\omega\Delta$ (look angle.) This variance is easier to see in Figures 4-9.
5. Figures 4-9 demonstrate that the bounds are effected by look angle to a much larger extent when ρ is increased to 0.4. As much as 5dB (Figs. 5,8) is observed at SNR = 0.1.
6. Comparing Figures in 4-9 with like SNR and M shows that different delays are effected quite differently as $\omega\Delta$ varies; i.e. CRMB (1,1), (2,2), (7,7) or (14,14) all vary differently with ρ and $\omega\Delta$.
7. Pursuing the question of how much clustering is effective when spatial noise is present, Figures 10 through 12 plot the variance bounds vs sensor number M. Figure 12 holds array length constant. We conclude that delay variances are reduced much less for M changing from 9 to 15 than for M changing from 3 to 9.

The last comment is meant to be one of the basic conclusions of this study: that for significant spatially correlated noise, there is a point beyond which it does not pay to increase sensor number in a cluster when the purpose is to reduce delay variance between clusters.

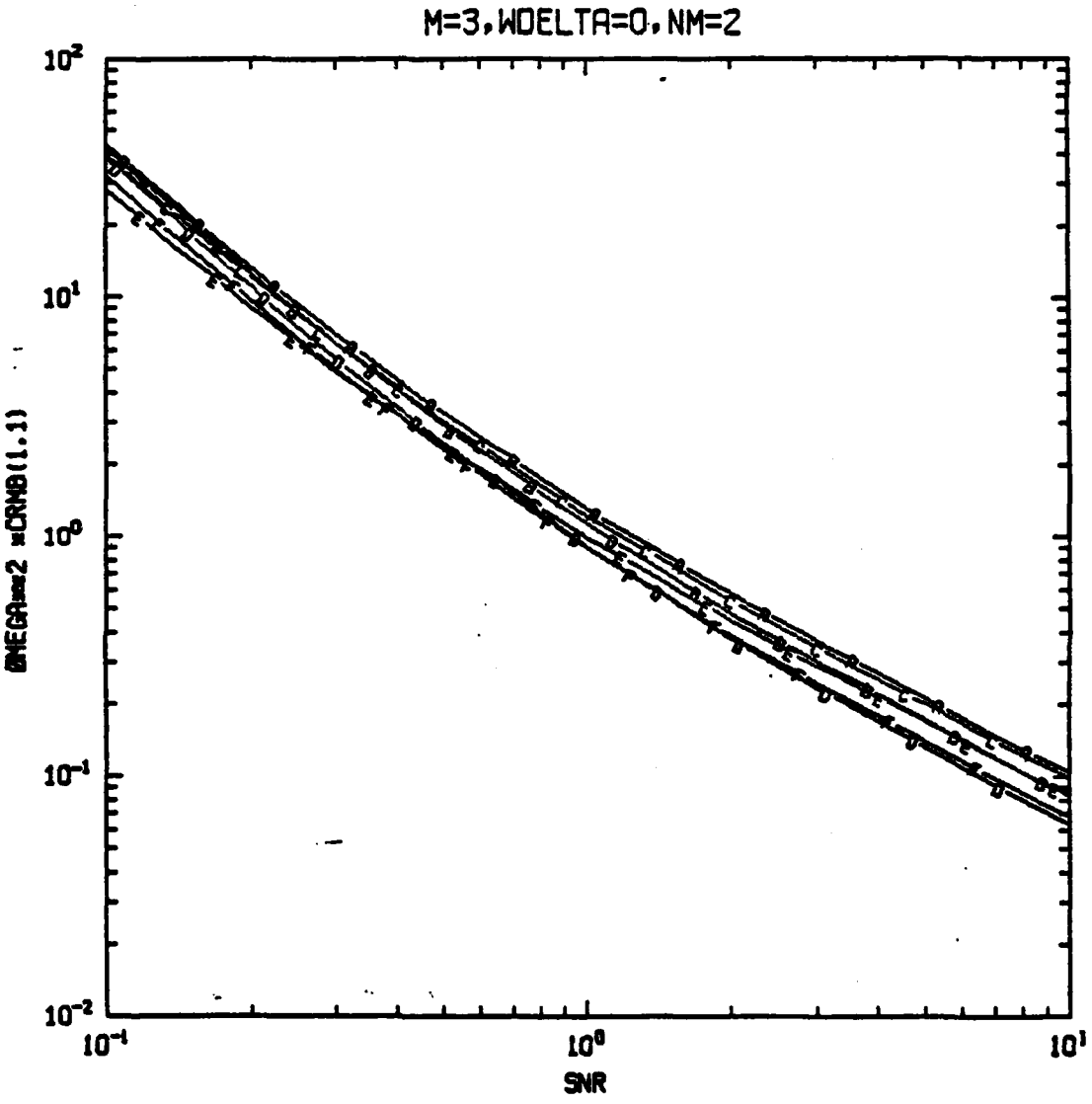


Figure 2. CRMB(1,1) vs SNR.M=3. A--ρ=0, B--ρ=0.2, C--ρ=j0.2; D--ρ=0.4, E--ρ=j0.4, F--ρ=0.4(1+j).ωΔ=0

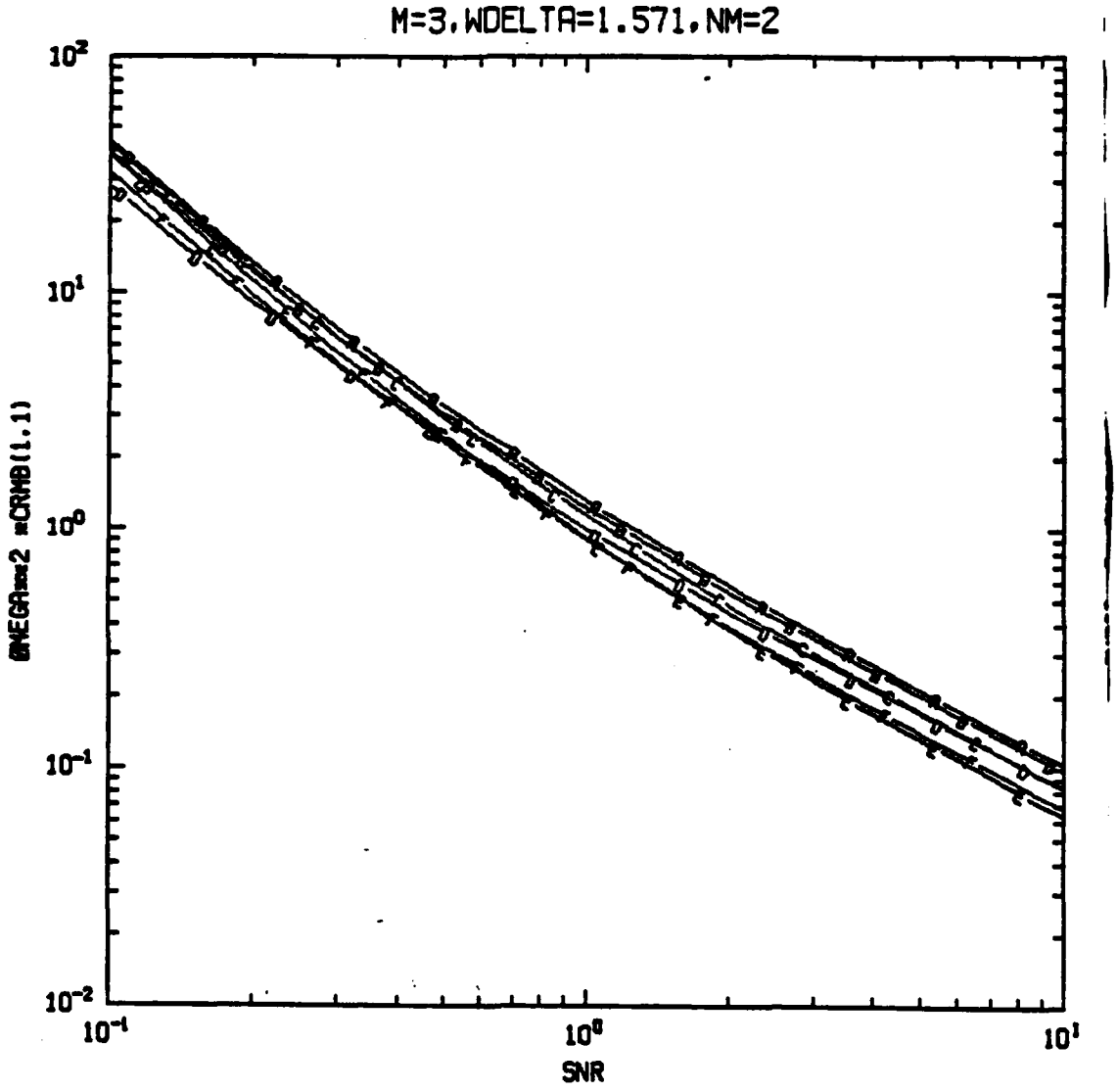


Figure 3. CRMB(1,1) vs SNR.M-3. A--ρ=0, B--ρ=0.2, C--ρ=j0.2; D--ρ=0.4, E--ρ=j0.4, F--ρ=0.4(1+j).$\omega\Delta = \pi/2$

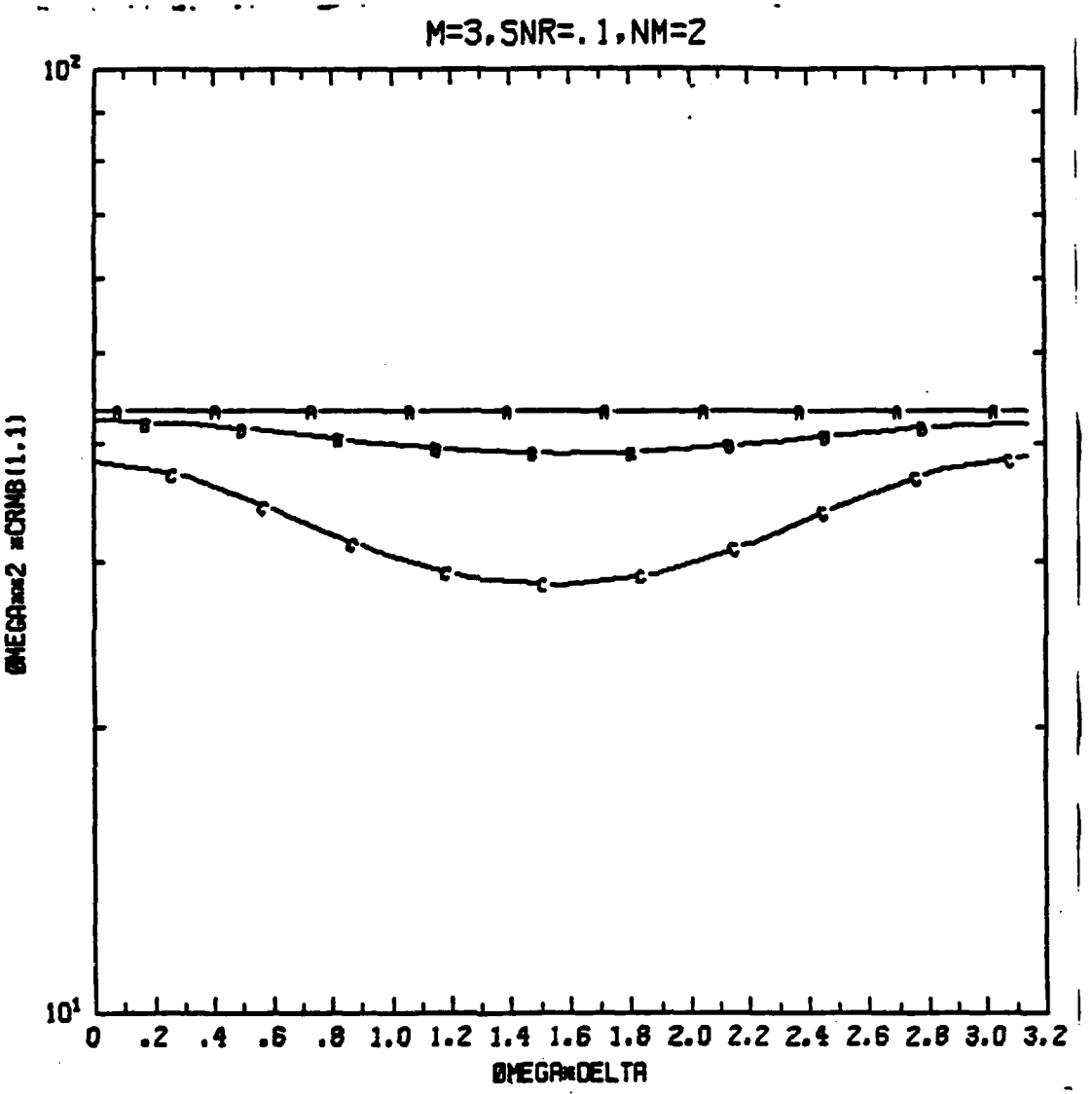


Figure 4. CRMB(1,1) vs $\omega\Delta$. M=3. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=0.1

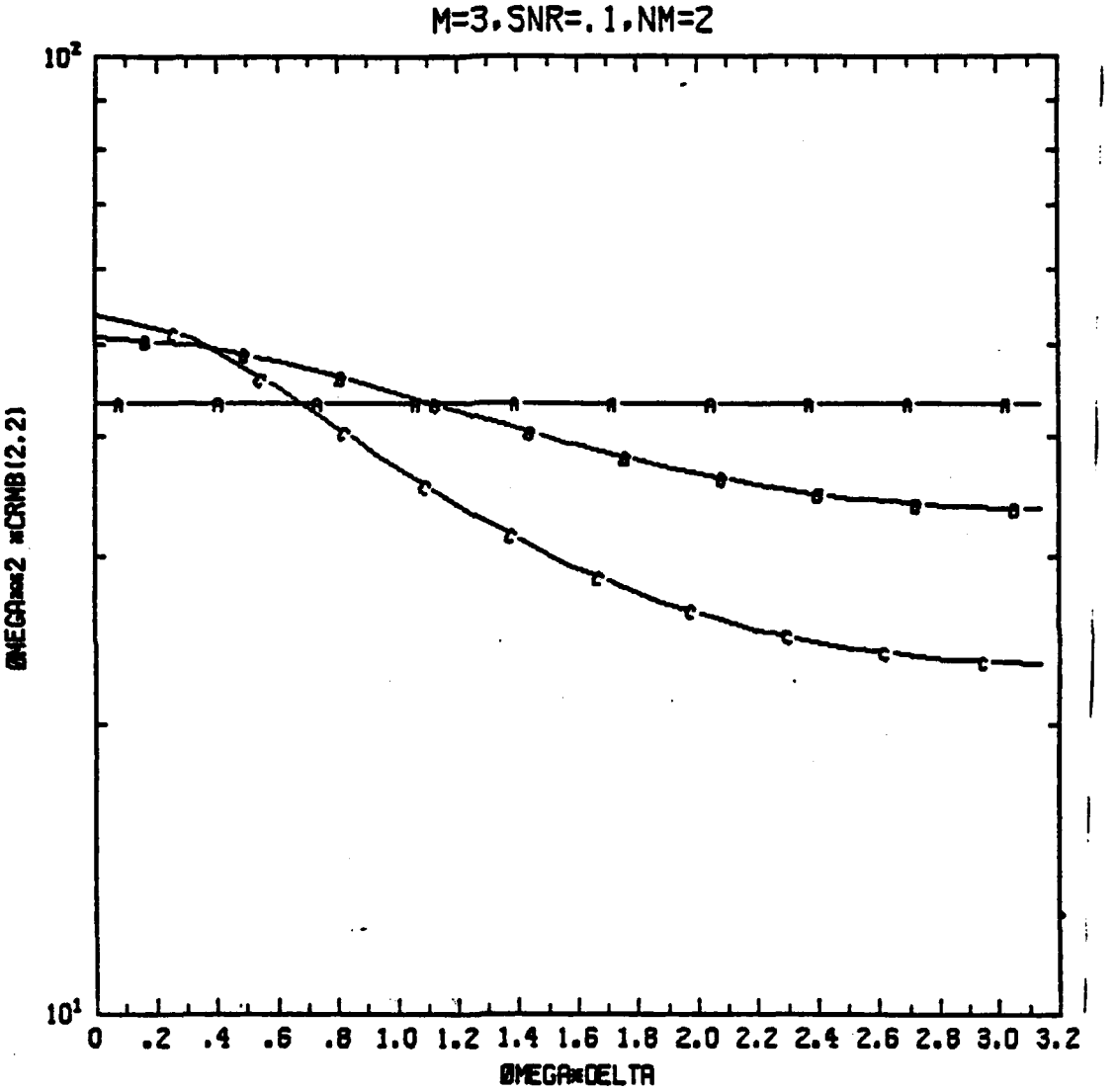


Figure 5. CRMB(2,2) vs $\omega\Delta$. M=3. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=0.1

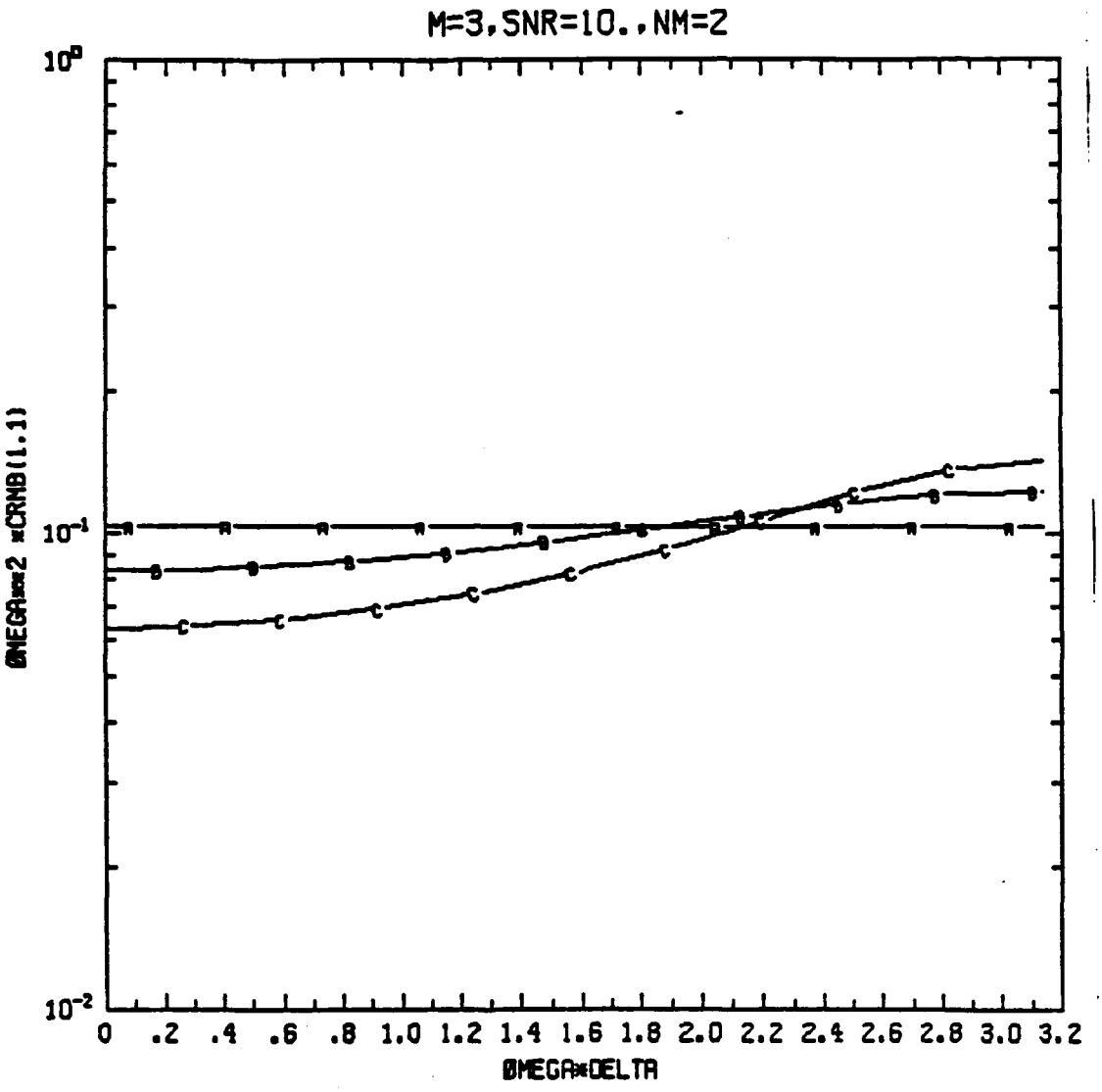


Figure 6. CRMB(1,1) vs $\omega\Delta$. M=3. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=10.0

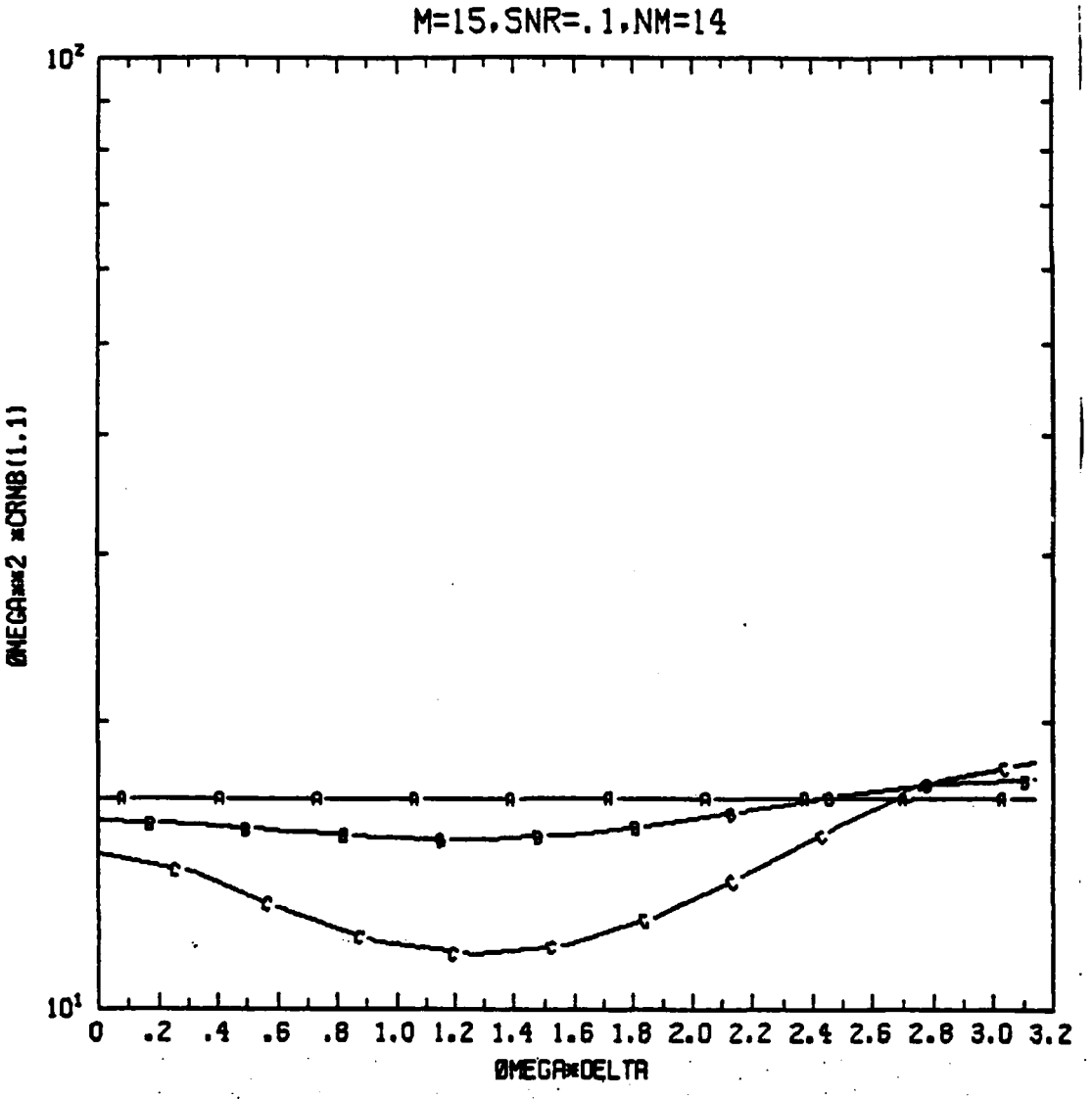


Figure 7. CRMB(1,1) vs $\omega\Delta$. M=15. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=0.1

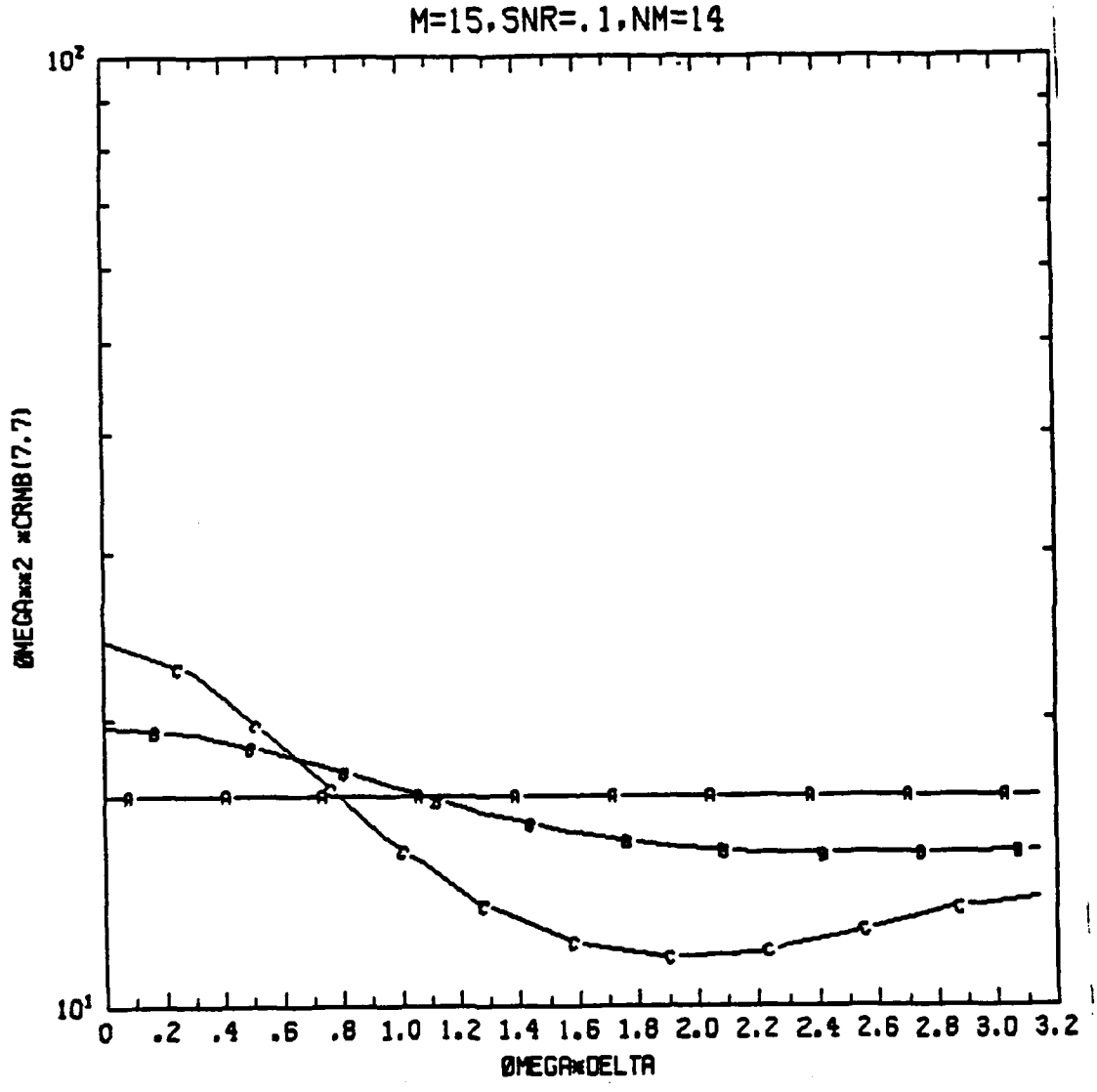


Figure 8. CRMB(7,7) vs $\omega\Delta$. M=15. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=0.1

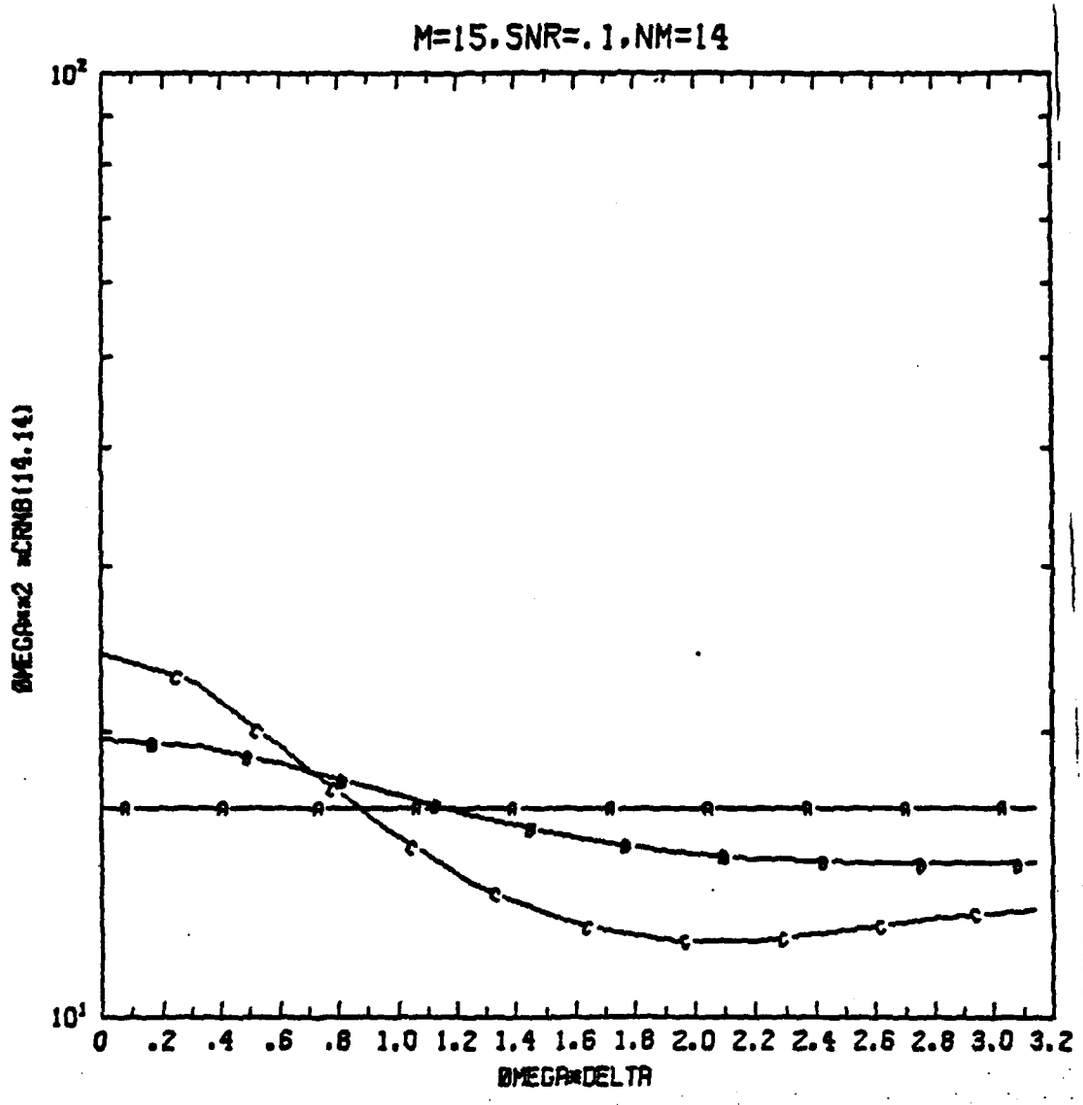


Figure 9. CRMB(14,14) vs $\omega\Delta$. M=15. A--ρ=0.0, B--ρ=0.2; C--ρ=0.4.SNR=0.1

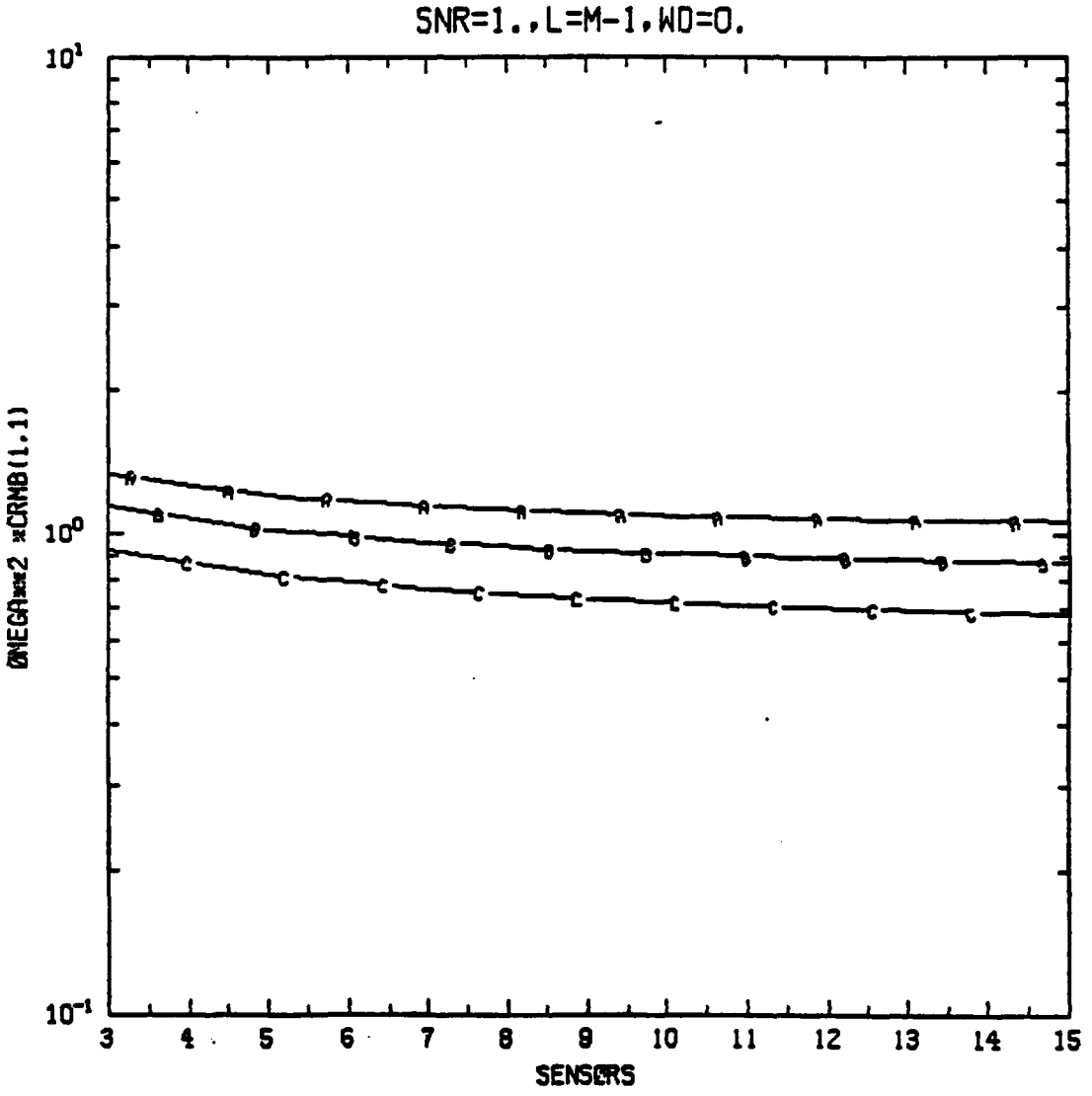


Figure 10. CRMB(1,1) vs M. SNR=1.0. A--ρ=0.0, B--ρ=0.2, C--ρ=0.4. Equally spaced sensors, full N-matrix.

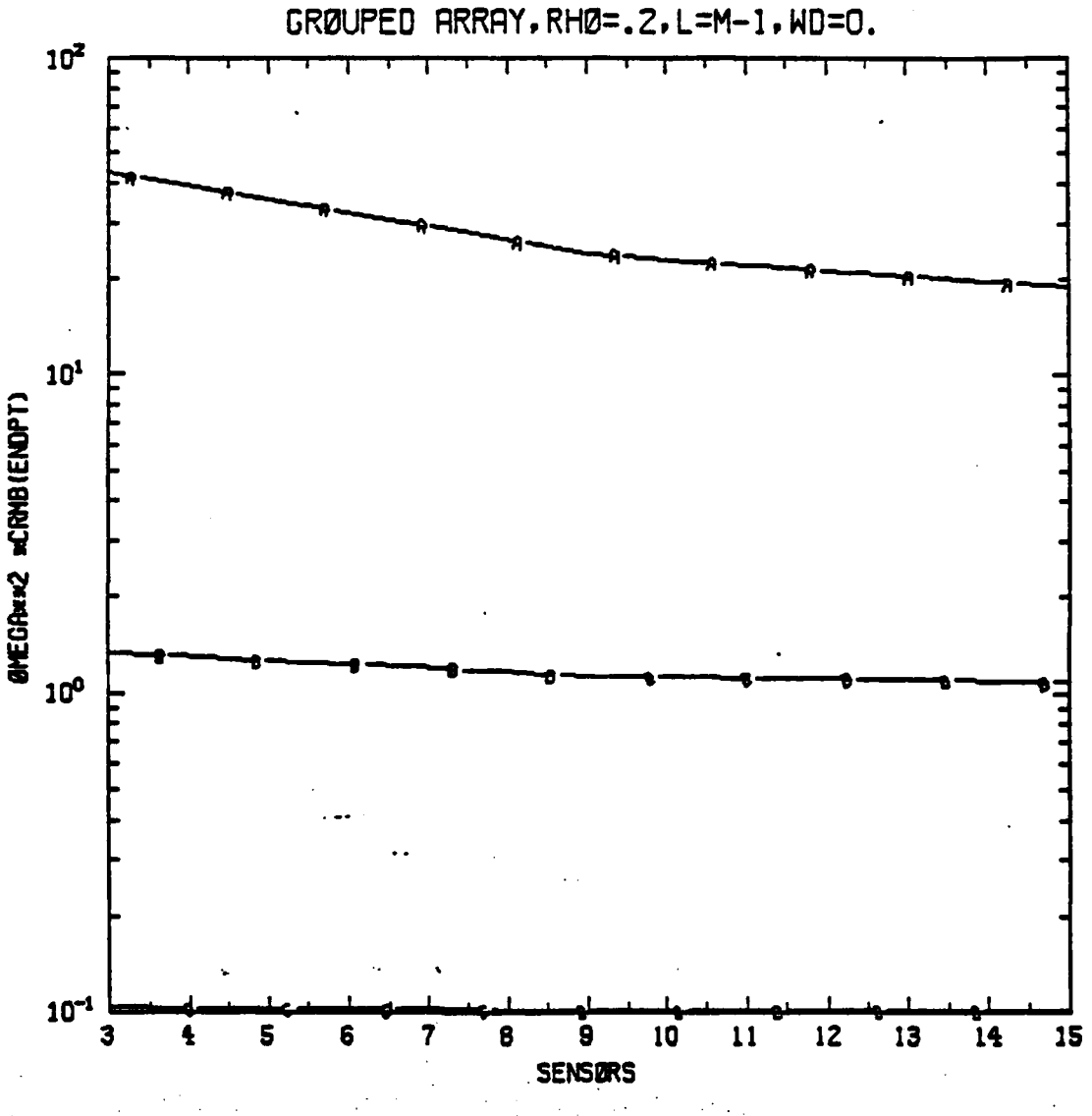


Figure 11. CRMB(M-1,M-1) vs M. ρ=0.2. A--SNR=0.1, B--SNR=1.0, C--SNR=10.0. M sensors have constant array length; clusters of 1, 3, and 5 elements at center and ends of array. No noise correlation between clusters.

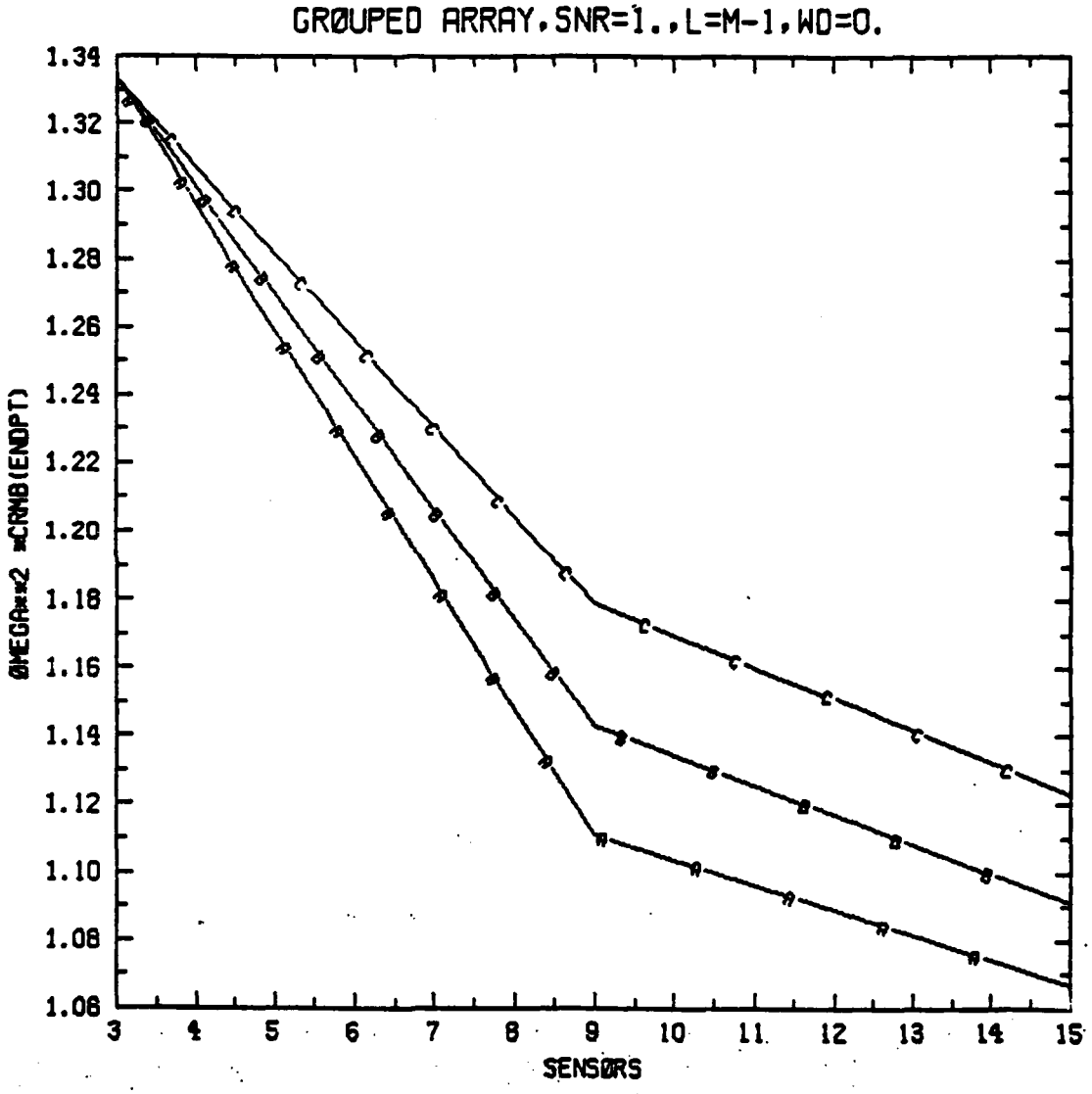


Figure 12. CRMB(M-1,M-1) vs M. SNR=1.0. A--ρ=0.0, B--ρ=0.2, C--ρ=0.4. M sensors have constant array length; clusters of 1, 3, and 5 elements at center and ends of array. No noise correlation between clusters.

REFERENCES

1. G. C. Carter, "Variance bounds for passively locating an acoustic source with a symmetric line array", J. Acoustic Soc. Am., Vol. 62, No. 4, Oct. 1977, pp. 922-926.

2. E. Weinstein, "Optimal source location and tracking from passive array measurements", IEEE Trans. A.S.S.P., Vol. ASSP-30, No. 1, Feb. 1982, pp. 69-76.

3. W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays", IEEE Trans. I. T., Vol. IT-19, No. 5, Sept. 1973, pp. 608-614.

4. W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation", J. Acoust. Soc. Am., Vol. 58, No. 1, July 1975, pp. 201-207.

5. P. M. Schultheiss, "Locating a passive source with array measurements, a summary of results", Proc. ICASSP, 1979, pp. 967-970. IEEE publication no. CH1379-7/79/0000-0967.

6. N. L. Owsley, "Time delay estimation in a sensor array", IEEE Trans. ASSP, Vol. ASSP-29, No. 3, June 1981, pp. 519-523.

7. B. F. Cron and C. H. Sherman, "Spatial-Correlation functions for various noise models", J. Acoustical Soc. Am., Vol. 34, No. 11, Nov. 1962, pp. 1732-1736.

8. A. G. Piersol, "Time delay estimation using phase data", IEEE Trans. ASSP, Vol. ASSP-29, No. 3, June 1981, pp. 471-477.

9. M. J. Hinich, "Frequency wavenumber array processing", J. Acoust. Soc. Am. Vol. 69, No. 3, Mar. 1981, pp. 732-737.

10. R. L. Kirlin, D. F. Moore, and R. F. Kubichek, "Improvement of delay measurements from sonar arrays via sequential state estimation", IEEE Trans. ASSP, Vol. ASSP-29, No. 3, June 1981, pp. 514-519.

11. N. L. Owsley and J. F. Fay, "Optimum turbulent boundary layer induced noise suppression with suboptimum realizations," J. A. S. A., Vol. 66, No. 5, Nov. 1979, pp. 1404-1411.

CHAPTER 3

A ROBUST RUNNING-WINDOW DETECTOR AND ESTIMATOR FOR STEP-SIGNALS IN CONTAMINATED GAUSSIAN NOISE

Abstract

An N-point window is applied to noisy data to recover stepped signals in non-Gaussian noise. Robust measures of signal step level and noise distribution spread are used to detect sequential clusters of data points which are statistically significantly different, thereby detecting the step. Using conventional analysis-of-variance methods, but with robust parameter estimates, false alarm probabilities are set reasonably accurately and miss probabilities and signal level estimates are shown by simulation to yield good results. Applications to Kalman filtering, seismic and well-log data, and image processing are indicated.

I. Introduction

The detection of steps of signal level in noisy data is a common problem. Adaptive Kalman filters estimate measurement bias and unknown step inputs to the process [1,2]. Image processing and segmentation often rely on detecting regions of constant level [3]. The robust median filtering method for extracting stepped signals in the non-Gaussian noise of well-logs has been found to be an improvement over such linear methods as linear prediction, Markov editing, and running average [4]. The median itself is a robust measure of location [5] which is resistant to values of wildpoints in the minority of the sample. Robust measures of spread are also appropriate when the error distribution is now known.

In the following we suggest robust methods for the detection of steps in non-gaussian noise. The noise used in the simulation is contaminated Gaussian, which is typical of heavy-tailed densities. Thus a step-detector of signals in such noise must not only be able to locate the step but estimate its level, while being resistant to the extreme values of the wild points in the sample window. Our methodology is related through its non-linear approach to those of median and order filtering [12-14].

We show by simulation that the suggested robust and resistant algorithm can be designed to give predictable results, giving stepped output with excellent step location accuracy even in low signal-to-noise ratios.

II. Running Cluster Detection

We wish to know if within the data window of length N there is a step in the noise-free signal. If there is a step we want to know its location and size. Several statistical tests are appropriate. Conventional tests on normal data are compared with robust methods.

In order to determine the most likely position of the step -- if it exists -- within the window, we cluster the j points X_i on the left in the window into cluster C_0 and the N - j points on the right into C_1. That j which maximizes the cluster-difference to spread-within-clusters ratio is assigned the step location point. Two robust methods are proposed for this purpose.

Robust cluster algorithm number 1

For any j, the cluster difference $\hat{S}$ is defined as the median difference [6, p. 133]

$$\hat{S} = \underset{\alpha,\beta}{\text{med}} \{X_{k-N+j+\beta} - X_{k-N+\alpha}\} , \qquad (1)$$

$$\alpha = 1, 2, \ldots , j \quad , \quad \beta = 1, 2, \ldots , N-j.$$

The within-clusters noise variance will be estimated with the robust formula

$$\hat{\sigma}_n^2 = \frac{1}{N-2} \frac{j(j-1)}{\Delta(\mu\alpha)} \sum_{\alpha}' ([X_{k-N+\alpha} - \text{med}\{X_{k-N+\alpha}\}]^2 (1 - \mu_\alpha^2)^4)$$

$$+ \frac{(N-j)(N-j-1)}{\Delta(\mu\beta)} \sum_{\beta}' ([X_{k-N+j+\beta} - \text{med}\{X_{k-N+j+\beta}\}]^2 (1 - \mu_\beta^2)^4) \qquad (2)$$

where

$$\Delta(\mu_\alpha) = [\sum'(1 - \mu_\alpha^2)(1 - 5\mu_\alpha^2)][-1 + \sum'(1 - \mu_\alpha^2)(1 - 5\mu_\alpha^2)] \qquad (3)$$

and $\Delta(\mu_\beta)$ is similar; α and β index over the members of clusters C_0 and C_1 respectively and $\sum'$ indicates only those members of the sum for which $\mu^2 \leq 1$, and where for example

$$\mu_\alpha = \frac{X_{k-N+\alpha} - \text{med}\{X_{k-N+\alpha}\}}{9\ \text{med}\{|X_{k-N+\alpha} - \text{med}\{X_{k-N+\alpha}\}|\}} \qquad (4)$$

Formula (2) combines robust variance estimation [5], with linear residual techniques [7]. Thus the μ_α, μ_β provide weights for the data within their respective clusters, the $\sum'$ over α and β include only those points in the sum-

of-squares which are not too wild, and the factors such as $j(j-1)/\Delta(\mu_\alpha)$ equal unity for small σ_n and no wild points, reducing $\hat{\sigma}_n^2$ to a typical variance estimate for the two-level, single factor problem in analysis of variance. The degrees of freedom (d.o.f.) for $\hat{\sigma}_n^2$ is suggested [5] to be 0.7 times the usual N - 3. The d.o.f. for $\hat{S}^2$ is taken to be 1, analogous to the sum-of-squares of two means around the population mean. The signal to noise ratio or variance ratio is estimated by

$$\hat{SNR} = \hat{S}^2/\hat{\sigma}_n^2 \tag{5}$$

This SNR is not the one which we wish to maximize over j. To optimize j, we must allow greater variation within cluster; therefore we just set $\mu_\alpha = \mu_\beta = 0$ for all α, β, in the calculation of $\hat{\sigma}_n^2$. Allowing for infinite SNR the algorithm must first test for zero in denominator expressions. Further we would not wish to assign just one (or maybe even just two) value to a cluster. That constrains $2 \leq j \leq N-2$.

In addition we must test the hypothesis that S = 0. Given the above d.o.f. we accept

$$H_0: \quad S = 0 \text{ if } F^* < F_{1,0.7(N-3)}(\gamma), \tag{6}$$

where we suggest

$$F^* = \frac{N(S/2)^2}{\sigma_n^2}, \tag{7}$$

and γ is cumulative area to the left of the F-density (with one and 0.7(N-3) d.o.f.) at the abscissa $F_{1,0.7(N-3)}(\gamma)$. For example if $\gamma = 0.50$ and N = 11, $F_{1,5.6}(0.50) \approx 0.520$. Thus $1 - \gamma = P_F$ = probability of false alarm if the data were purely gaussian.

Robust cluster algorithm number 2

This algorithm uses the conventional ANOVA method, but the data is pre-processed with a three or five-point median filter (a length small with respect

to the window). Typical results show that both $\hat{S}$ and $\hat{\sigma}_n^2$ have been reduced because of the rounding effect of the median at the step location. Justusson (in [8]) gives the variance for such a filtered normal white sequency (M = 3) to be $0.4487\sigma^2$ [8]. He also shows the rounding effect at the step location.

The d.o.f. measure is often taken to be 2TB where 2B is the two-sided equivalent signal bandwidth and T is the data record length [9,10]. We can determine 2B from a known discrete auto-correlation function $R_{\tilde{x}}(\tau)$ of the median-filtered sequence $\tilde{x}_i$, assuming Δ seconds between samples. The equivalent band-width B is

$$2B = \int_{-\infty}^{\infty} S_{\tilde{x}}(f)df / S_{\tilde{x}}(0)$$

$$= R_{\tilde{x}}(0) / \int_{-\infty}^{\infty} R_{\tilde{x}}(\tau) \qquad (8)$$

Discretizing these measures the d.o.f. in measuring $\hat{\sigma}_n^2$ with N samples is

$$\nu_e = N \; R_{\tilde{x}}(0) / \sum_{k=-M+1}^{M-1} R_{\tilde{x}}(k) \qquad (9)$$

From Justusson [8] we have for M = 3, $R_x(0) = 0.4487$, $R_x(1) = 0.2480$, and $R_x(2) = 0.1177$. Because we remove means from the C_0 and C_1 squares, the overall d.o.f. for N points over the two clusters is $\nu_e - 2$. This gives the clusters respectively

$$\nu_0 = j \; R_{\tilde{x}}(0) / \sum R_{\tilde{x}}(k) \qquad (10a)$$

and

$$\nu_1 = (N - j) R_{\tilde{x}}(0) / \sum R_{\tilde{x}}(k) \qquad (10b)$$

d.o.f., and the variate estimate in the general case is

$$\hat{\sigma}_n^2 = \frac{1}{\nu_e - 2} [(\nu_0 - 1) \hat{\sigma}_0^2 + (\nu_1 - 1) \hat{\sigma}_1^2] \qquad (11)$$

where $\hat{\sigma}_i^2$ are variances for C_0 and C_1 calculated with $\nu_i - 1$ degrees of freedom.

We suggest that using standard ANOVA applied to pre-median filtered data with $\hat{\sigma}_n^2$ calculated as in (10) and (11) is appropriate for detecting and estimating a step. However, because of the effect that pre-median filtering has on the data near the step location, the signal estimate is smaller than with other methods. This may lead to erroneous acceptance of H_0: $S = 0$ for small step-to-noise ratios, where other methods would not err. In addition the d.o.f. estimate per Equations (9-11) is only useful to the extent that $R_{\underset{\sim}{x}}$ is known. We will see in our results that these formulas gave P_F lower (and thus P_M = probability of miss higher) than designed.

III. Simulations for F-test Error Probabilities for Steps in Gaussian and Non-Gaussian Noise

Data has been produced to allow evaluation of both robust clustering algorithms. The programs yield maximum SNR clustering within the running data window. However, for determining merit factor (defined later) and probabilities of miss and false alarm the window samples are independent. The number of window-length data sets is 375. When steps are present they occur $3(N-1)/2 + 1$ points apart.

As discussed following Equation 5, to optimize the clustering SNR a more conventional SNR is used; this is implemented by setting the data weights equal to unity. We still have a robust SNR measure in that cluster medians rather than means are used in estimating the steps. The weights are retained (Equations 3-4) for $\hat{SNR}$ which is used in the F-statistic, Equation (7), which is to be compared to $F_{1,0.7(N-3)}(\gamma)$. If $F^* < F$, then no step is declared; otherwise a step is declared. The decision creates possibilities for both miss and false alarm, which probabilities are estimated by the simulations.

The type of contaminated Gaussian noise considered is such that the density of x(k) (with no signal step) is

$$f(x) = (1 - \varepsilon)\, N_1(0,1) + \varepsilon N_2(0,\sigma_2^2) \ . \tag{12}$$

Thus the contamination is also Gaussian, but x has heavier tails when $\sigma_2 > 1$.

In addition a merit factor calculation is evaluated. This figure of merit is discussed by Pratt [3] with regard to detection of edge pixels in images. The formula is

$$\text{fom} = \frac{1}{I_N} \sum_{i=1}^{I_A} \frac{1}{1 + \alpha d^2} \ , \tag{13}$$

where $I_N = \max(I_I, I_A)$ and I_I and I_A are respectively true number and actual detected number of steps in the test data, α is a scaling constant, and d is the location error distance count between a true step and one detected. We have used $\alpha = 1/9$; for this value fom becomes 0.9 if for example all steps are detected but mislocated by one point. If on the other hand there were no false alarms, and all detections were accurately located, but 20% were missed, then fom = 0.8. At the other extreme, if fraction p of the detections are false alarms ($I_N = \frac{1}{1-p} I_I$), and those p are all d counts away from a true step, then fom = $\frac{1+\alpha d^2(1-p)}{1+\alpha d^2}$ (this only applies when there are true steps (SNR > 0) in the data). It can be seen that

$$\text{fom} \to \begin{cases} (1 + \alpha d^2)^{-1}, & p \to 1 \\ 1 - p \qquad\quad , & d^2 \to \infty \end{cases} \tag{14}$$

Note that in our figuring, a step detected with $d > 2$ is considered both a false alarm and a miss, but that $d \leq 2$ constitutes a valid detection. Now consider the results of Figures 1-3 wherein $\varepsilon = 0$ (uncontaminated Gaussian).

Figure 1 shows the probability of miss vs. SNR for N = 15. The curves within the Figure are for γ (the variable in $F(\gamma)$) equal to 0.5, 0.9, and 0.99. These curves show that at SNR $\to$ 0, Pr(miss) $\simeq \gamma$, which is to be expected because as S $\to$ 0 with fixed noise the density of the data is essentially that of the noise. At higher SNR's Pr(miss) decreases for fixed γ and increasing N.

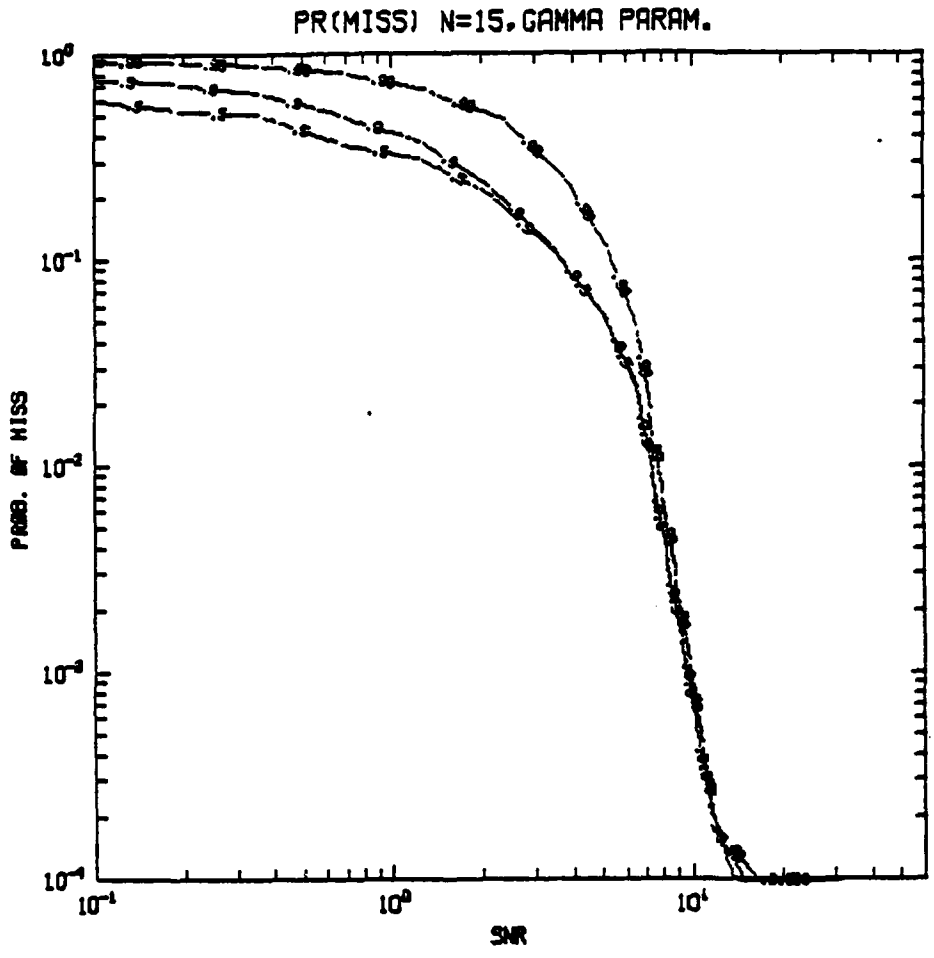


Figure 1. Probability of miss vs. SNR, N = 15, no contamination, algorithm number 1.

Figure 2 tends to verify the use of the d.o.f. prescribed to $\hat{\sigma}_n^2$ in the F-value [5], i.e., 0.7(N-3). However, some inaccuracy is noted. The curves should be level at the value $1 - \gamma$; that is, Pr(false) = $1 - \gamma$. The curves are generally too low, even though they are somewhat better for larger N. For small N, such as N = 5, undoubtedly the variance estimates are pessimistically too large, reducing the Pr(false). But for N larger (9, 15)) and γ large (.9, .99), the d.o.f. approximation seems to be fair.

Figure 3 is the figure-of-merit curve for N = 15. We found that the trend is toward higher merit for N larger (as reasonable). Also as N increases there is a lower SNR at which point higher γ are more meritorious than lower γ. We realize that as SNR $\rightarrow$ 0, γ = 0.5 will maximize fom and minimize probability of either error type, but at some higher SNR maximation is via larger γ. Similar figures may be found in [11] for N = 5 and 9.

We now discuss and compare results when noise is contaminated Gaussian. Specifically the noise is ε = 0.05 contaminated at scale 7.14 (σ_2^2 = 51.0) such that σ_x^2 = 3.5.

Figure 4 is comparable to Figure 1. With contamination the probability of miss at any SNR is greater. The contamination undoubtedly raises the noise variance estimates, lowering F* values below the decision threshold, and causes the miss. It has been shown in [11] that conventional noise variance estimators will be much larger than the robust estimators, decreasing F* even more and increasing the probability of miss.

Curves for false alarm and fom have also been run with noise contamination. Very little difference is noted. Again this is likely due to the fact that the robust estimators make it difficult to let wild points alter the signal estimate while at the same time increasing somewhat the noise estimate. Thus the trend is to slightly reduce, if anything, the false alarm rate as

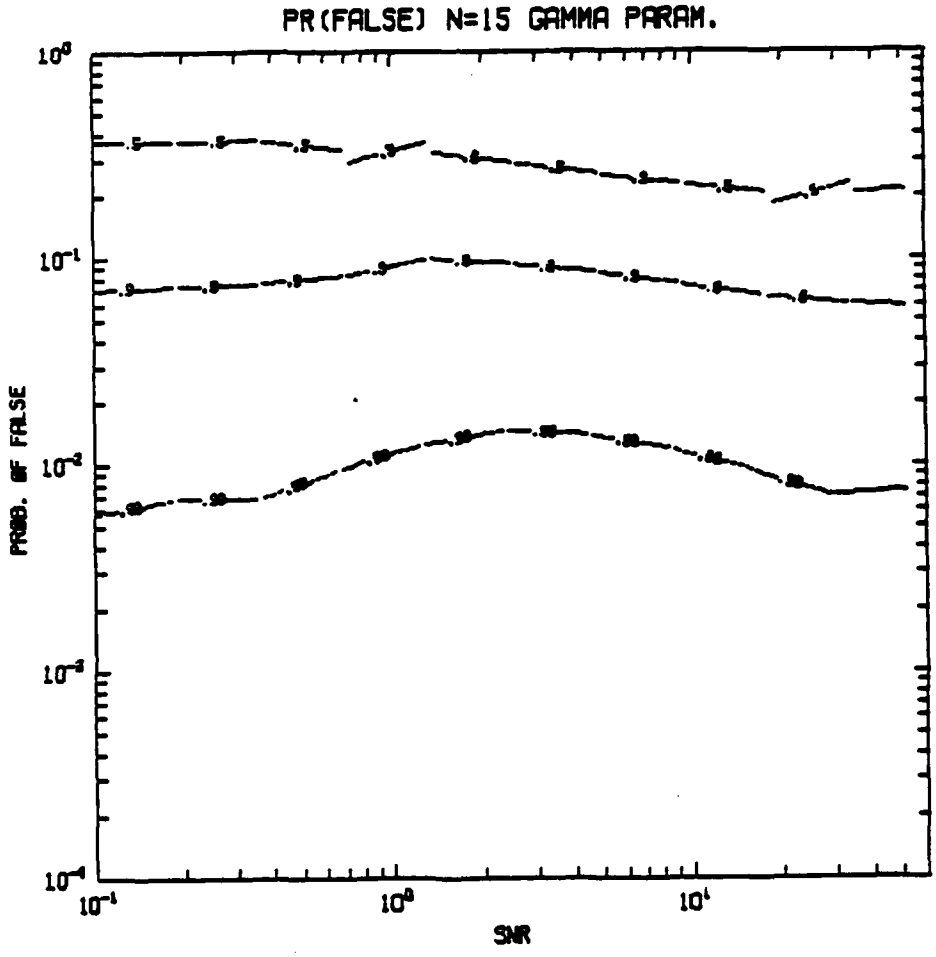


Figure 2. Probability of false alarm vs. SNR, N = 15, no contamination, algorithm number 1.

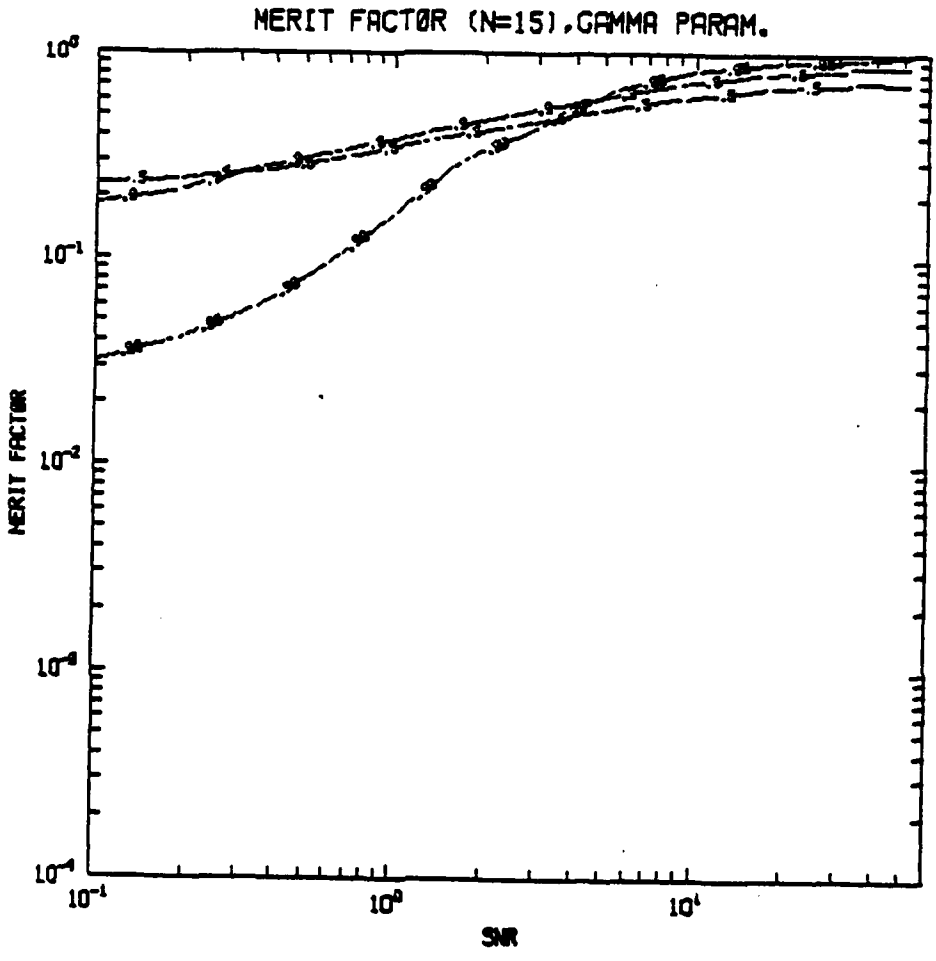


Figure 3. Figure of Merit vs. SNR, N = 15, no contamination, algorithm number 1.

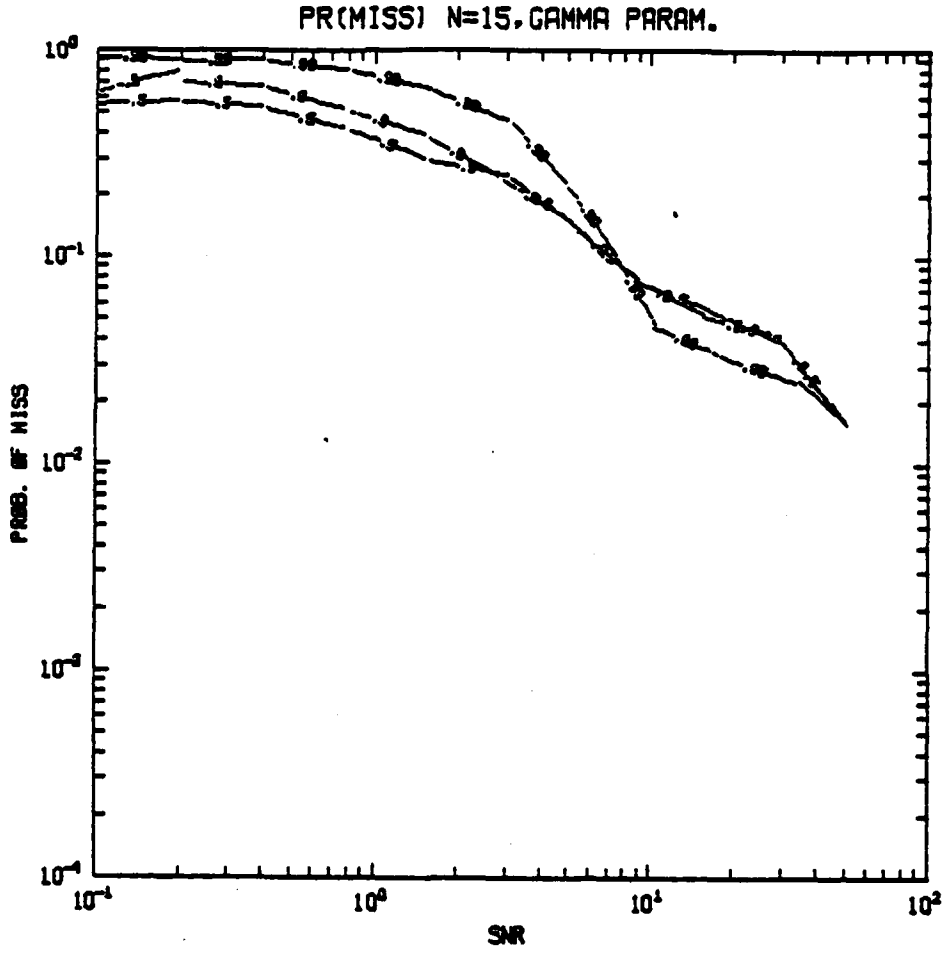


Figure 4. Probability of miss vs. SNR, N = 15, contamination ϵ = .05, σ_2 = 7.14, algorithm number 1.

contamination increases. This is perhaps a beneficial result of using the robust estimators of cluster centers and noise variances. Figure of merit does slightly decrease at larger SNR, due to increased miss rate.

We next tested robust algorithm number 2 with a pre-median filter length M = 3. Results are shown in Figures 5-7 for no contamination and in Figure 8 for the same contamination as earlier.

Comparing Figures 1 and 5, probability of miss with $\varepsilon = 0$ are higher for algorithm number 2 than those of algorithm number 1. An explanation is that pre-median filtering considerably smoothes to the middle of the step edge, reducing the measured spread between clusters.

Figure 6 for probability of false alarm, $\varepsilon = 0$, N = 15, algorithm number 2, should be compared to Figure 2 of algorithm number 1. It may be seen that false alarm rate has dropped considerably due to the pruning of wild points by the pre-median filtering. The cost however was the increased probability of miss, just discussed. An inappropriate estimate of d.o.f. is another explanation, causing the threshold to be too high.

IV. Simulations of Signal Estimation

In order to better visualize the results of using the running window clustering algorithm, example runs with different SNR and F-threshold were simulated. The noise process in each of these runs is the same. Its density is

$$p_n(n) = 0.95\ N(0,1) + 0.05\ N(0,51)$$

and $\sigma_n^2 = 3.5$. Signal steps occur at sample numbers 22, 44, 88, ... Window size is held fixed at N = 15, and SNR = S^2/σ_n^2, where S = true step size between clusters.

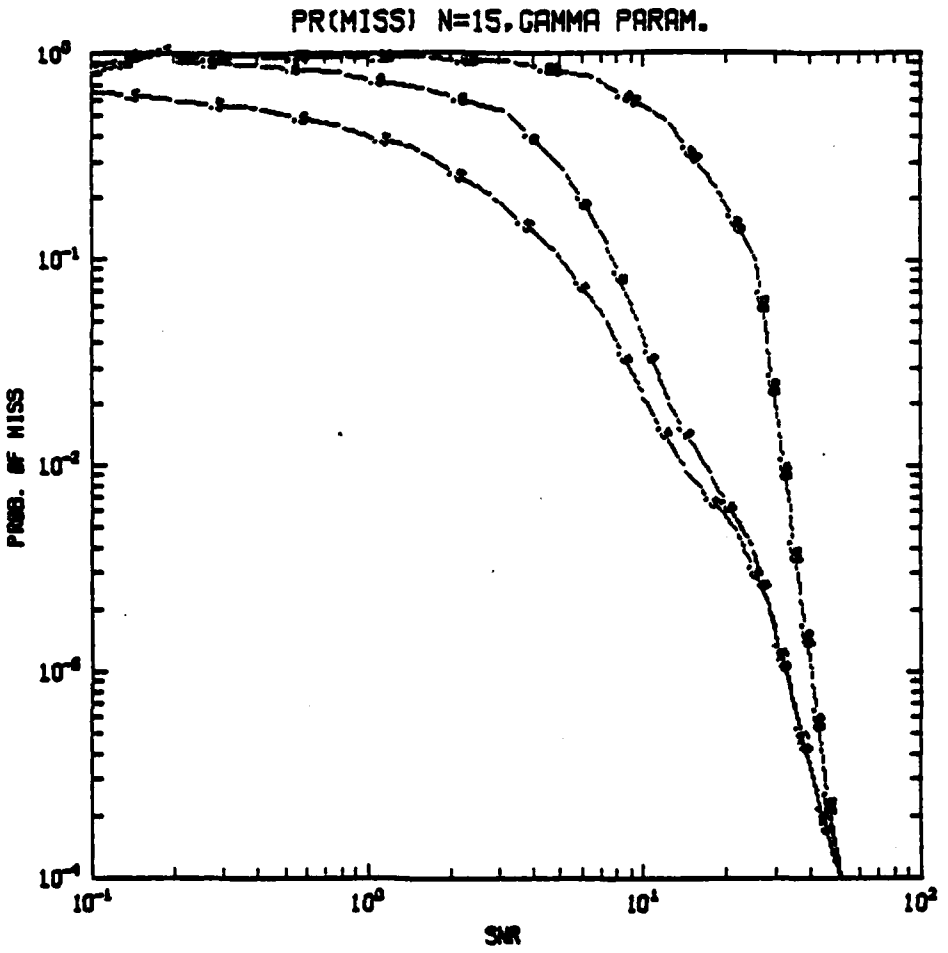


Figure 5. Probability of miss, no contamination, algorithm number 2, N = 15.

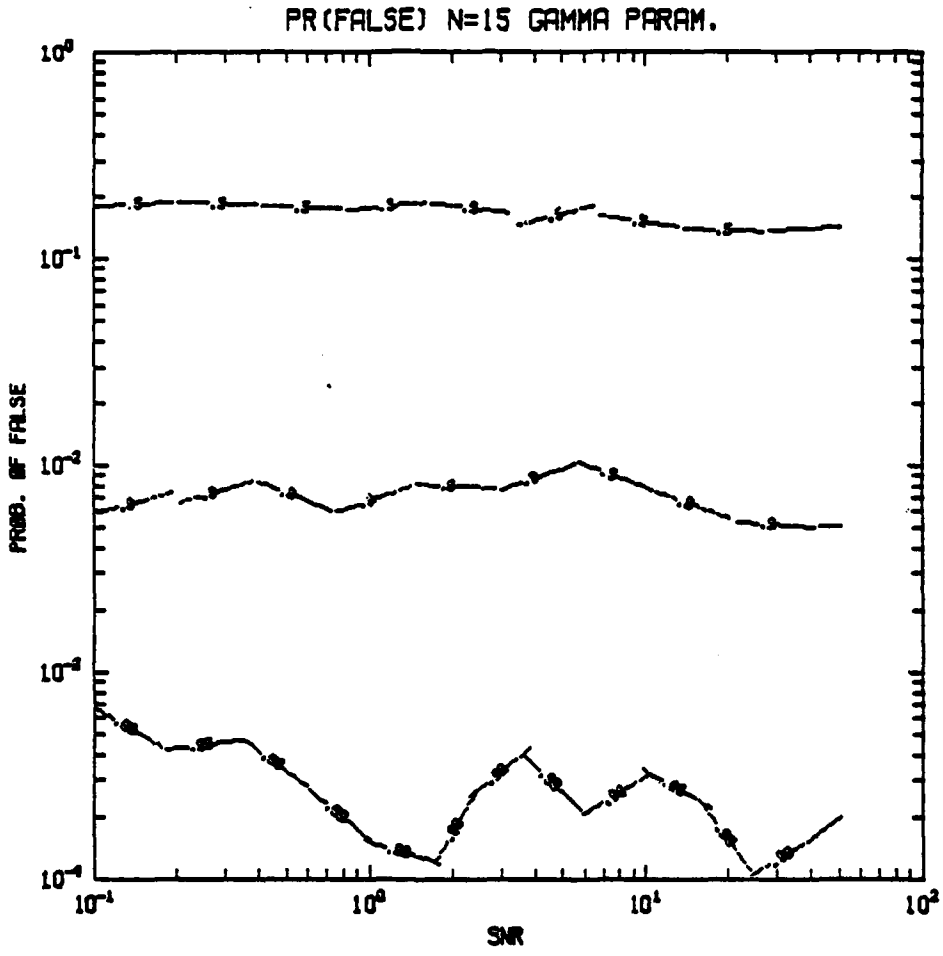


Figure 6. Probability of false alarm, $\varepsilon = 0$, algorithm number 2, N = 15.

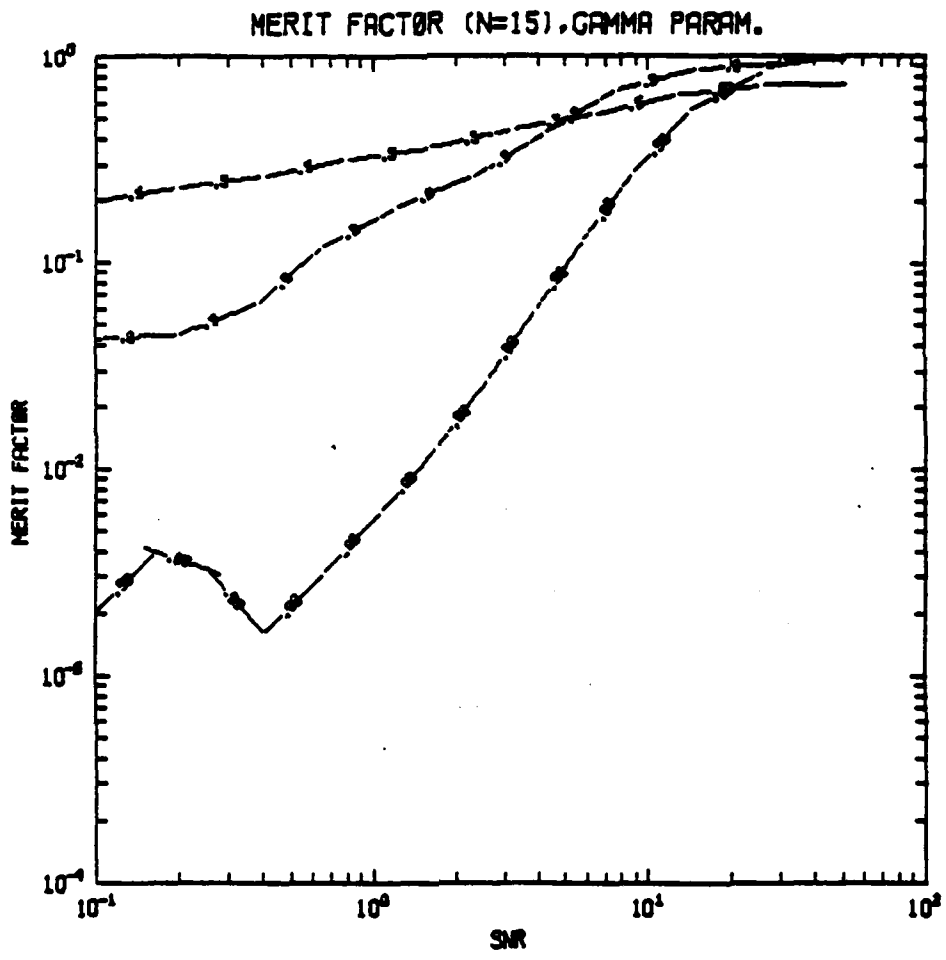


Figure 7. Figure of merit, $\varepsilon = 0$, algorithm number 2, $N = 15$.

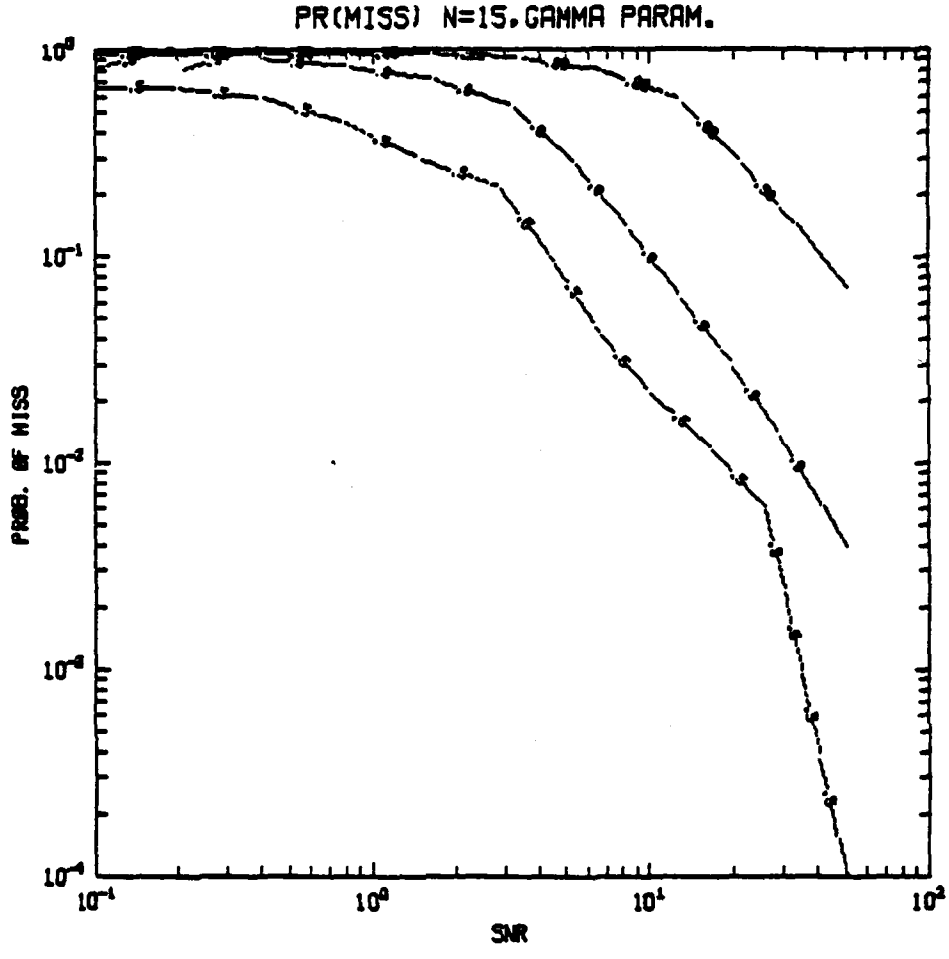


Figure 8. Probability of miss, $\epsilon = 0.05$, algorithm number 2, N = 15.

Two step-size estimators are used. Method 1 is as we have described before; the new level assigned to the output is the median of that cluster, C1, which contains the newest data. Method 2 does not assign an output level until the next step is detected; then it assigns the median of all samples between step locations. Nevertheless, step detection with method 2 uses only data within the window. This is necessary if the F-value is to remain fixed. Method 1 is necessary for real time, finite-delay decisions and signal level estimation as we find in Kalman filtering.

Figures 9-11 are for method 1 with SNR = 4, 2, and 1 respectively, and Figures 21-23 similarly show results of method 2. Method 2 should be better due to the larger number of samples possible between step detections than that in cluster (0 within the window). Figures 9-11(a) show the result for $\gamma = 0.5$ and Figures 9-11(b) show the result for $\gamma = 0.95$, or, respectively, false alarm thresholds 0.5 and 0.05. On the left is the input data and on the right are true and estimated signal plots. Figures 12-14 are for $\gamma = 0.95$, Method 2.

These figures verify expectations. Fewer steps are allowed with $\gamma = 0.95$, the step locator is surprizingly accurate even in considerable noise, and estimated step value is reasonably good and immune to wild points.

Comparing Figures 9-11(a) with Figures 12-14, it is not clear that method 2 is advantageous. More analytical work or simulation needs to be done.

V. Conclusions

This work has introduced robust statistical methods as a solution to the problem of detecting stepped signals in non-Gaussian noise. The method uses F-statistics and standard analysis-of-variance techniques for detecting steps, but the cluster means are medians and the noise variance estimates use

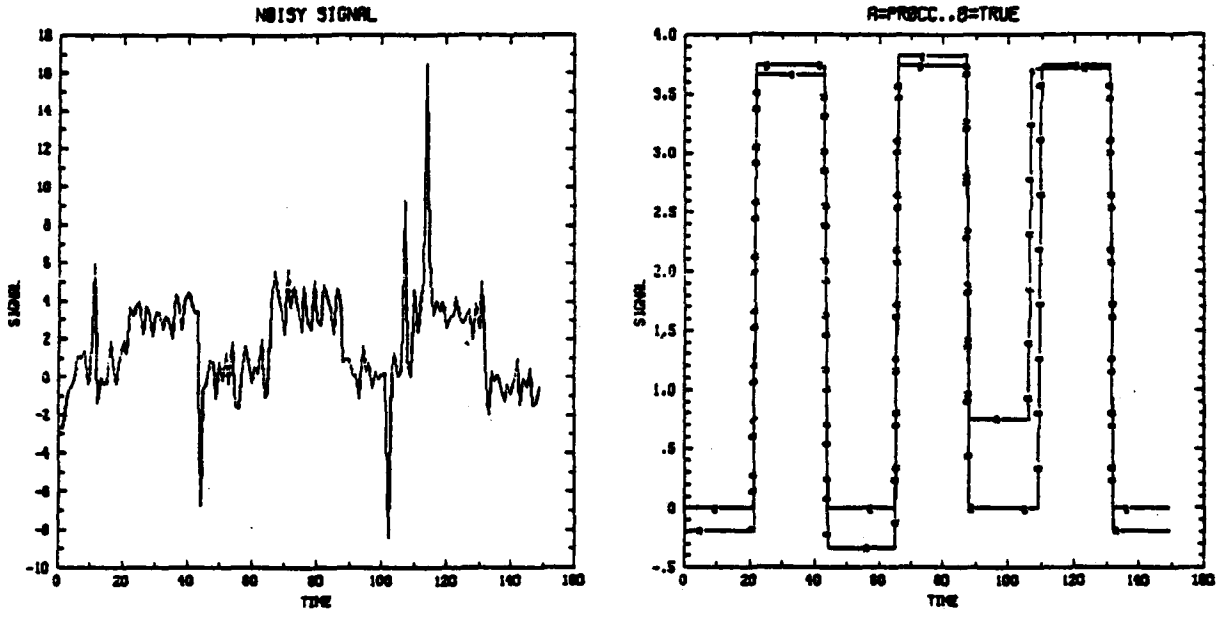


Figure 9-a. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 4, P_F = 0.5, N = 15.

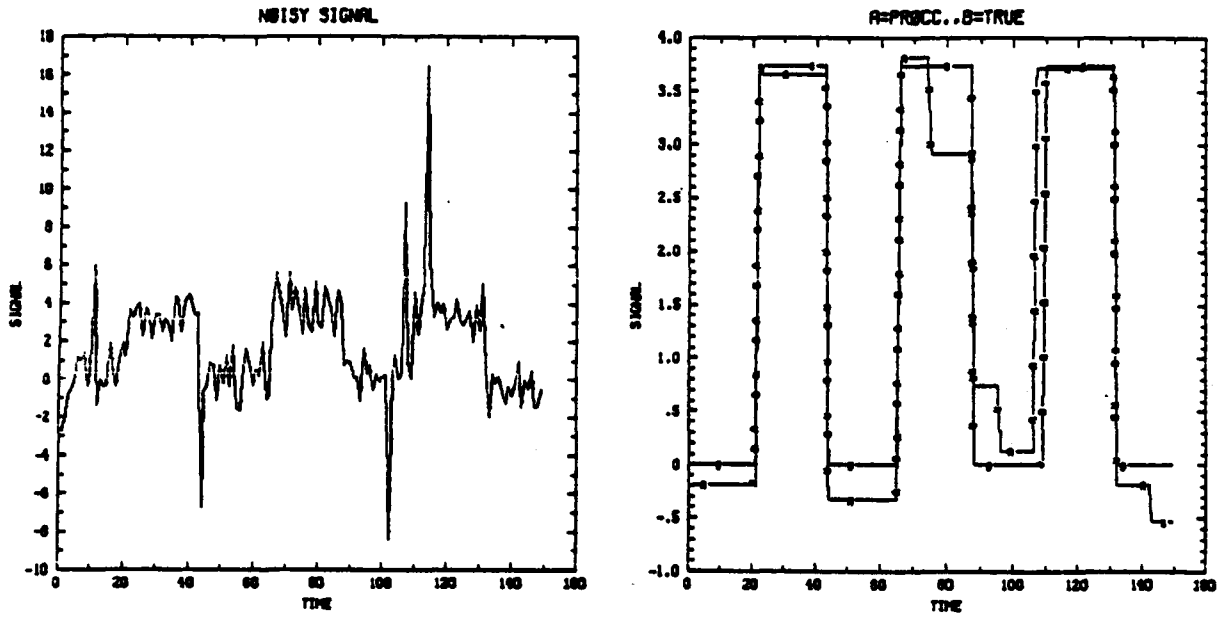


Figure 9-b. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 4, P_F = 0.05, N = 15.

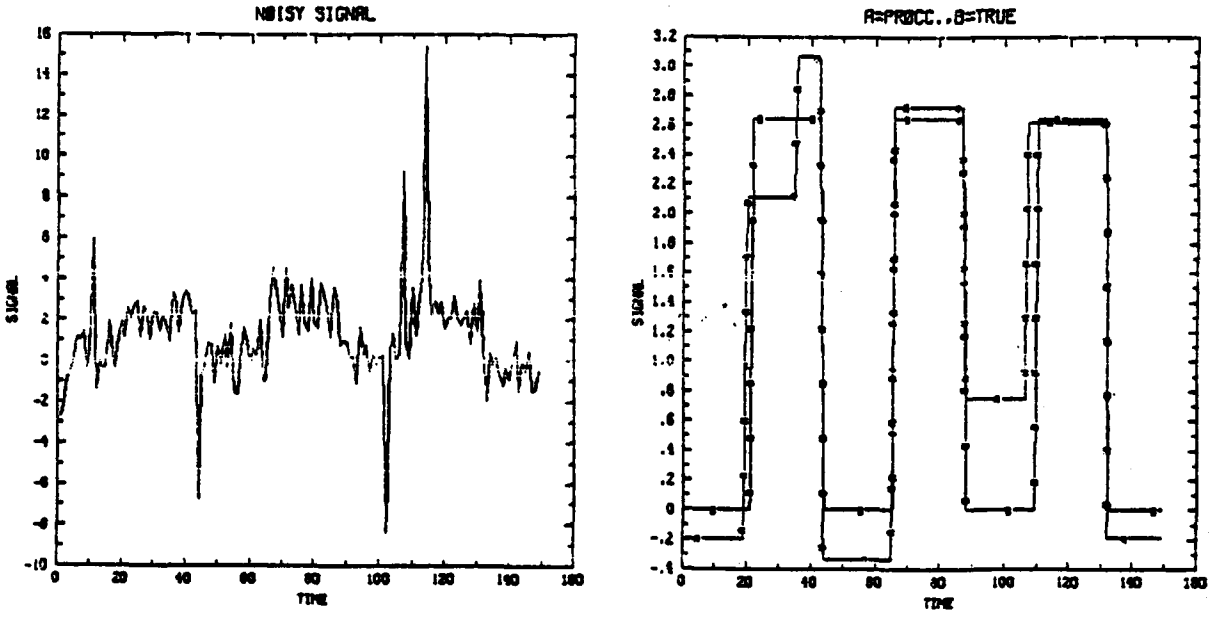


Figure 10-a. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 2, P_F = 0.5, N = 15.

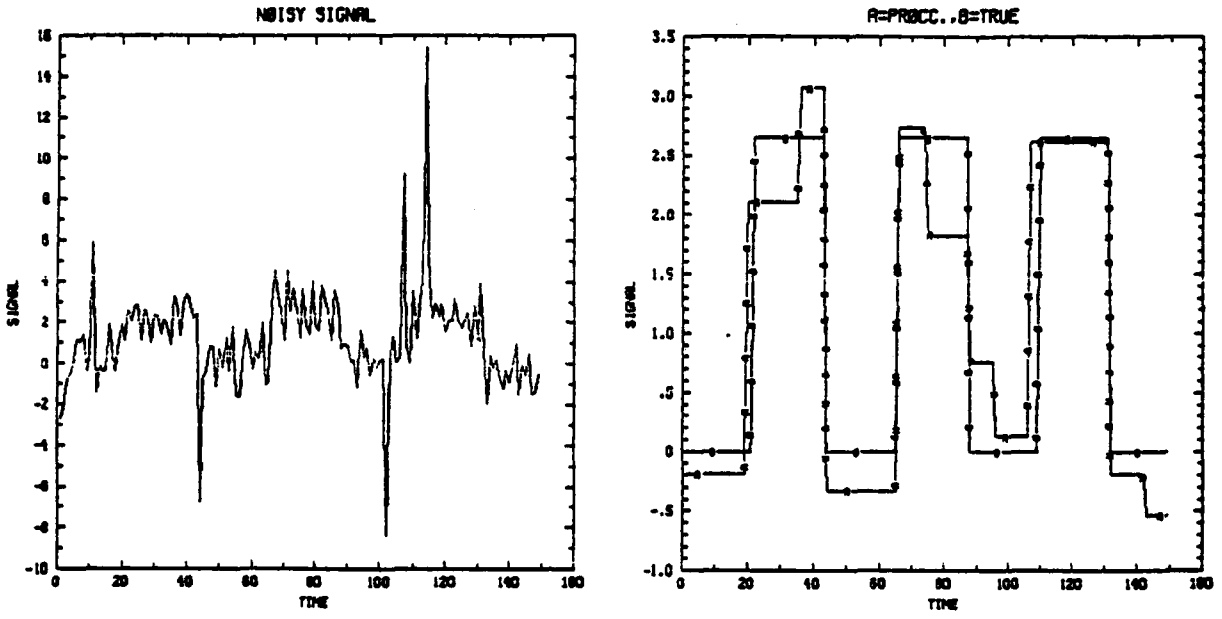


Figure 10-b. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 2, P_F = 0.05, N = 15.

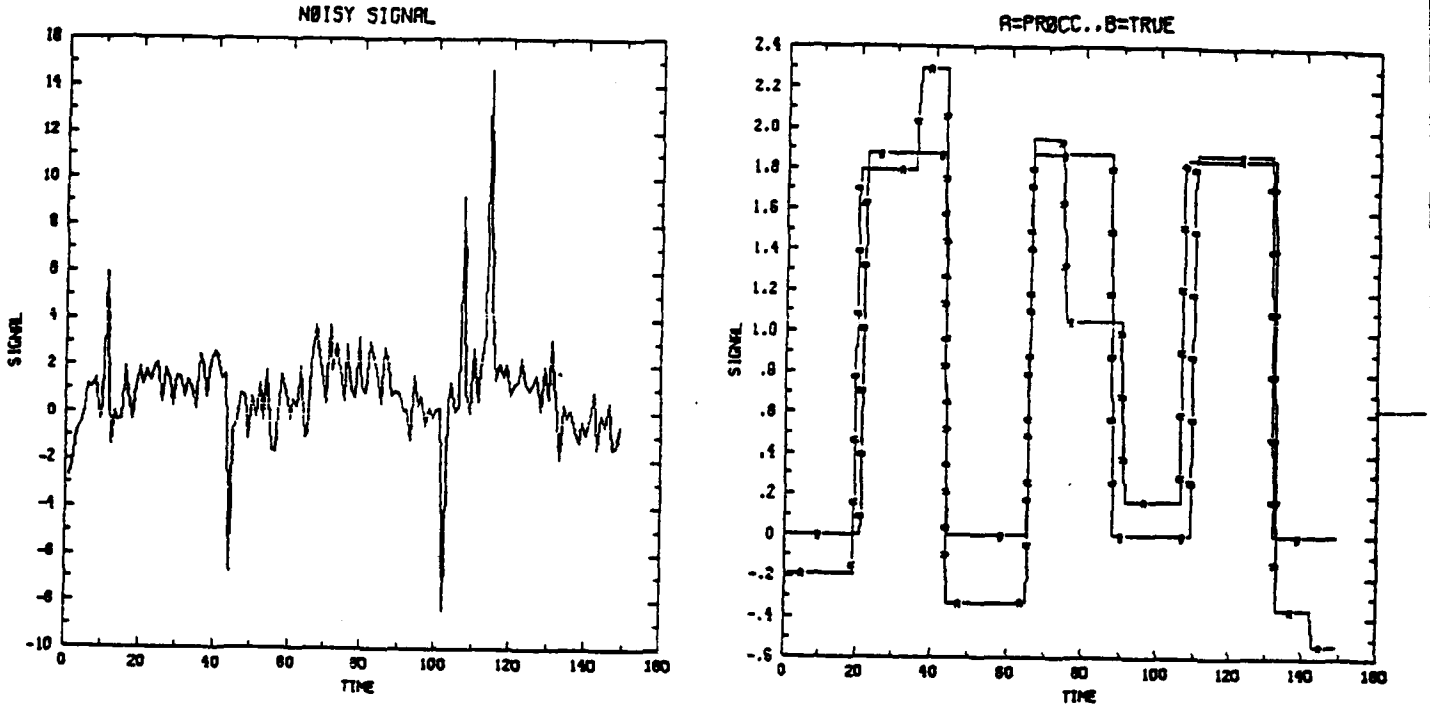


Figure 11-a. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 1, P_F = 0.5, N = 15.

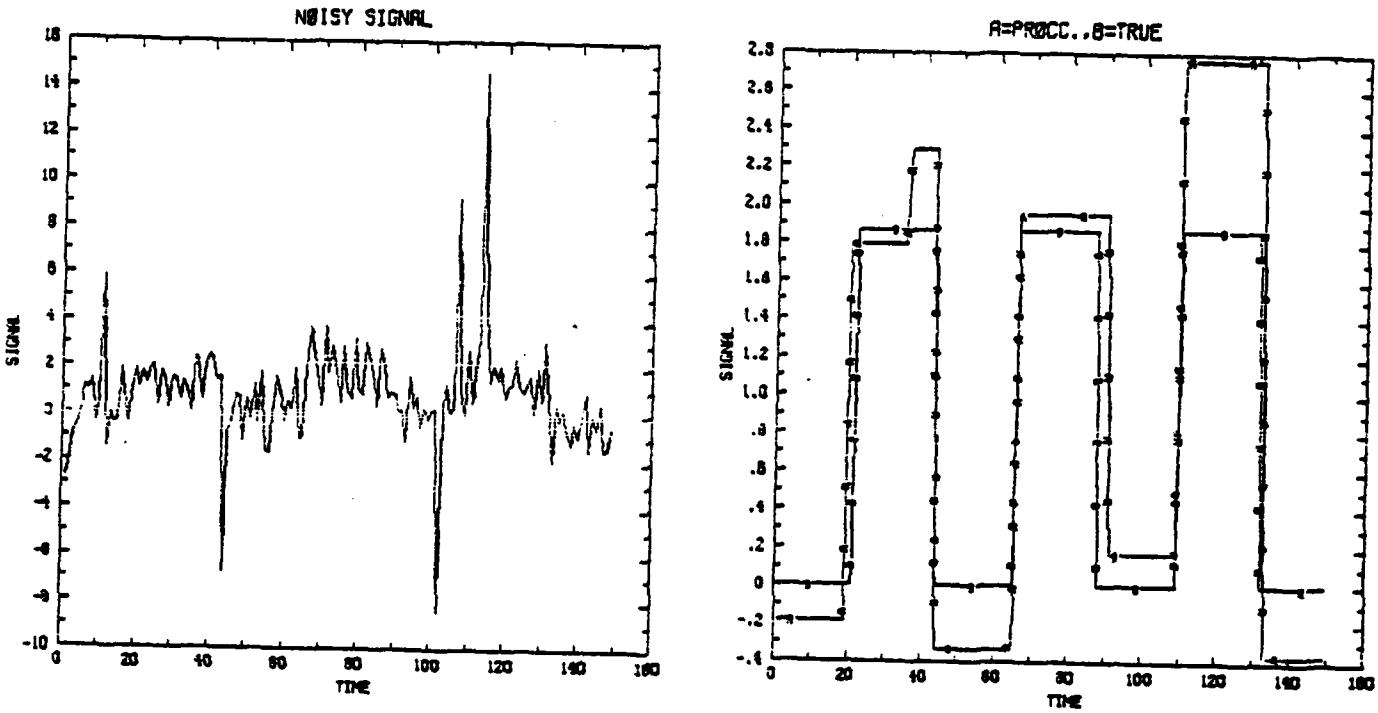


Figure 11-b. Method 1, true (B) and estimated (A) signal in noise from data (left). SNR = 1, P_F = 0.05, N = 15.

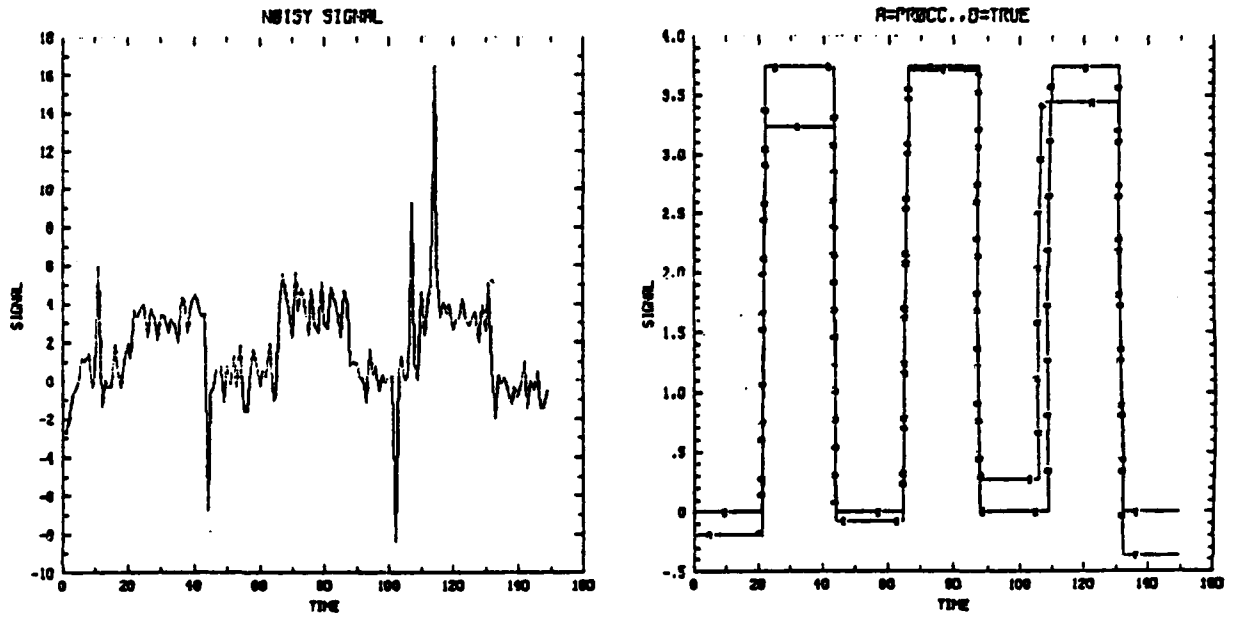


Figure 12. True (B) and estimated (A) signal in noise from data (left). SNR = 4, P_F = 0.05, N = 15.

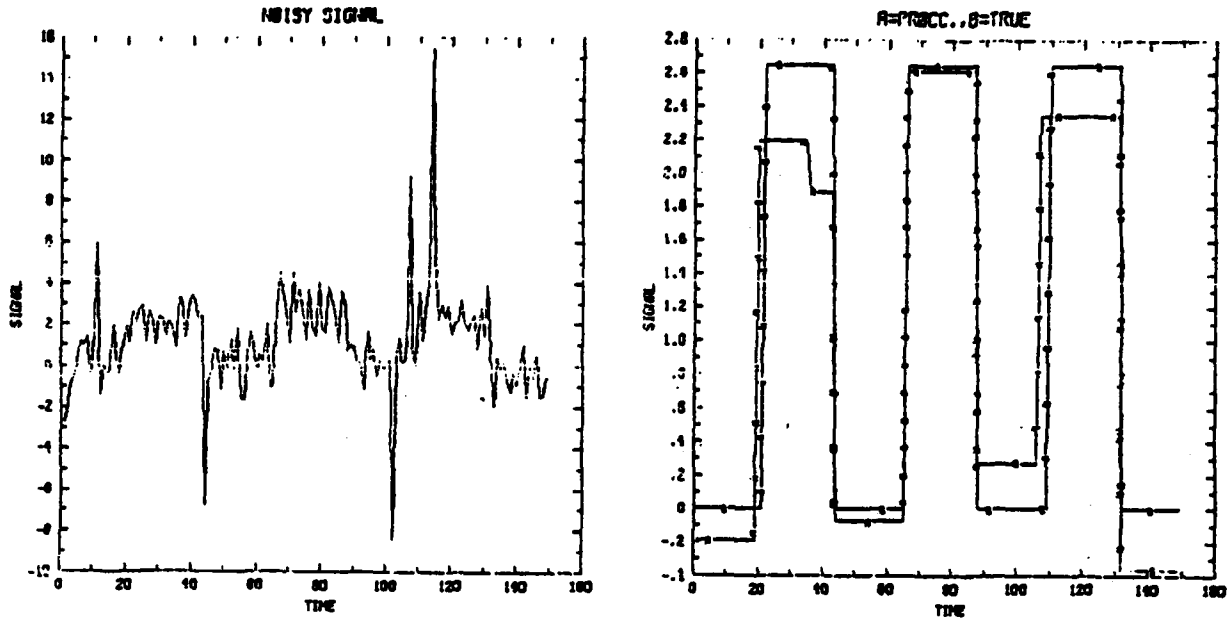


Figure 13. True (B) and estimated (A) signal in noise from data (left). SNR = 2, P_F = 0.05, N = 15.

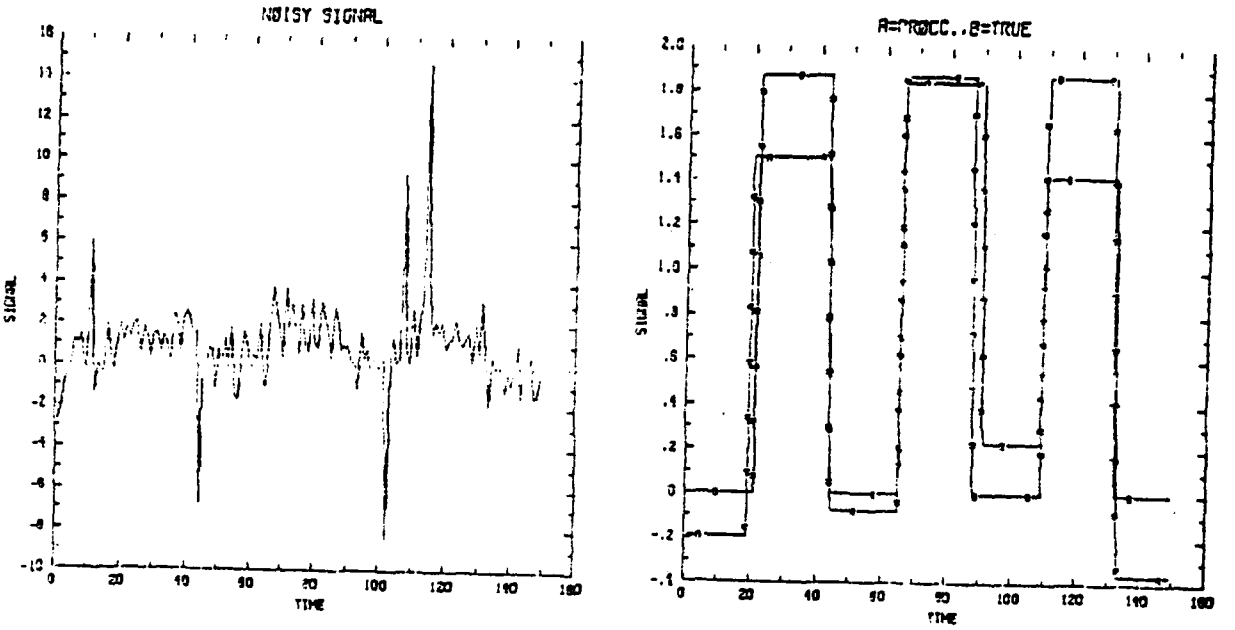


Figure 14. [illegible] (B) and estimated (A) signal in noise from data (left). SNR = 1, P_F = 0.05, N = 15.

a biweight method for robustness. It is shown that probability of false alarm can be closely predicted and that step locations and values are closely estimated. The algorithm in one of its variations should find uses in any field where constant valued data is to be located in heavy-tailed noise.

REFERENCES

1. Myers, K.A. and B.D. Tapley, "Adaptive Sequential Estimation with Unknown Noise Statistics," IEEE Transactions Automatic Control, vol. AC-21, No. 4, pp. 520-523. August 1976.

2. Groutage, F.D., "State Variable Estimation using Adaptive Kalman Filter with Robust Smoothing," Proceedings 22nd IEEE Conference on Decision and Control, pp. December 1983. IEEE no. 0191-2216/83/0000-0304.

3. Pratt, W.K., Digital Image Processing, Wiley, New York, 1978.

4. Bednar, J.B., "Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data," Geophysics, Vol. 48, No. 12, December 1983, pp. 1593-1610.

5. Mosteller, F. and J.W. Tukey, Data Analysis and Regression, Addison-Wesley, Reading, MA. 1977.

6. Hollander, M. and D.A. Wolfe, Nonparametric Statistical Methods, John Wiley and Sons, New York, 1973.

7. Neter, J. and W. Wasserman, Applied Statistical Models, Richard D. Irwin, Inc., London, 1974.

8. Huang, T.S. (editor), Two Dimensional Digital Signal Processing II, Transforms and Median Filters, Springer-Verlag, N.Y., 1981.

9. Blackman, R.B. and J.W. Tukey, The Measurement of Power Spectra, Dover, 1959.

10. Jenkins, G.M., and D.G. Watts, Spectral Analysis, Holden Day, San Francisco, 1968.

11. Kirlin, R.L. 1984 Annual Report, Project #N0014-82-K-0048, Office of Naval Research.

12. Bovik, A.C., T.S. Huang and D.C. Munson Jr., "A Generalization of Median Filtering Using Linear Combinations of Order Statistics," IEEE Transactions ASSP, vol. ASSP-31, No. 6, pp. 1342-1350. December 1983.

13. David, H.A., Order Statistics, Wiley, New York. 1980.

14. Gallagher, N.C. Jr., and G.L. Wise, "A Theoretical Analysis of the Properties of Median Filters," IEEE Transactions ASSP, vol. ASSP-29, No. 6, pp. 1136-1141. December 1981.

CHAPTER 4

ROBUST ADAPTIVE KALMAN FILTERING FOR SYSTEMS WITH UNKNOWN STEP INPUTS AND NON-GAUSSIAN MEASUREMENT ERRORS

Abstract

Target tracking with Kalman filters is hampered by target maneuvering and unknown process and measurement noises. We show that moving data windows may be used to analyze state and measurement error sequences, determining robust estimates of bias and covariance. For steps in the system forcing functions and non-Gaussian measurement errors the robust estimators yield improvements over linear bias and covariance estimators. Extensive simulations compare conventional, linear adaptive and robust adaptive average step responses of a first-order system filter. Quantities examined are state estimate, state error, process and measurement covariance estimates, Kalman gain and input step estimate.

I. INTRODUCTION

The tracking of targets is a difficult problem in practical situations. Measurements of target location are usually derived from time delay estimates, which is a subject in itself, and a large body of literature dealing with that alone has been compiled in recent years and is typified in a recent special issue [1]. The translation of time delays to target location and motion parameters is via nonlinear functions, and error variances on the delays result in confidence intervals on the location parameters [2, 3]. Hassab and Boucher have shown experimentally, however, that delay estimates are not Gaussian [4]. In many other filtering problems the measurements are also non-Gaussian. Several Kalman tracking schemes are examined by Hassab et al. [5], wherein the plant model is linear in the states while the measurement relationship is nonlinear in the states.

Previous efforts of applying Kalman filter methodology to tracking have been very useful, yet the practical problems of non-Gaussian noise and maneuvering targets have either been ignored or only lightly touched upon with the linearized approaches until recently. It may be reasonable to model target system process noise as Gaussian, but it is probably not reasonable to model either the deterministic (but unknown) system forcing function or the time-delay measurement noise as Gaussian. Recent papers by Bradley and Kirlin [6] and Weiss and Weinstein [7], as well as other previous papers giving simulations of time delay estimation algorithms, have pointed out the need for editing the delay measurements, particularly in low SNR or narrow band signal situations.

Forcing functions for the target motion system may be taken into account by the Kalman tracker by increasing the variance of the process noise, preferably adaptively. An example of a similar scheme is given by

Godiwala [8], wherein a bank of Kalman filters, each assuming a different mean process input, are adaptively weighted. A number of target maneuvers are simulated to test the algorithm. Results are good, although some limitations are discussed, and not too drastic velocity changes, termed semi-Markov in the simulation, have a 0.01 probability of occurring.

A more difficult and typical tracking problem has been addressed by Groutage [9, 10]. Based on work by Myers and Tapley [11], Groutage uses mean and covariance estimators from moving windows on the state and measurement error sequences to adapt to non-zero, deterministic acceleration inputs. He also employs a non-optimal Kalman gain modification and robust post-smoothing of the state estimates. These methods yield improved results over the linear adaptive methods in [11].

The non-Gaussian noise problem is also addressed by Boncelet and Dickinson [12], who suggest formulating the Kalman filter as a regression problem and performing robust regression. They show no results, however. Another approach without results is discussed by Agee and Turner (in [13]). Masreliez and Martin [14] develop an optimal Kalman gain to edit wild measurements with a specific non-Gaussian noise, but they do not consider step inputs. Meyr and Spies [15] use probability of outliers to edit wild points in the Kalman filter, but again do not adaptively deal with step inputs.

We propose tracking methods similar in kind to those of Groutage and Myers and Tapley, but adaptive robust estimators of location and spread are used in place of their linear estimators of mean and variance. The location estimator chosen is the median, which is known to precisely track noise-free step functions without the smearing effect inherent with running sample

means [16-17]. The robust covariance measure uses the biweight proposed by Tukey [18] and has good promise for dealing with the non-Gaussian delay estimates.

II. CONVENTIONAL AND ADAPTIVE KALMAN FILTERS

The system state equation is

$$x(k+1) = Ax(k) + u(k) + w(k), \tag{1}$$

where x is the state vector, A is a matrix, u(k) is an acceleration vector, and w(k) is a vector of white noises which account for either or both source and sensor array random motions.

The measurement vector is

$$z(k) = Hx(k) + v(k), \tag{2}$$

where v is an additive error vector, and H is a matrix.

The conventional estimator and error covariance update equations are [19]

$$\hat{x}(k) = A\hat{x}(k-1) + K(k)(z(k) - HAx(k-1)) \tag{3}$$

$$P(k) = (I - K(k)H)(AP(k-1)A^T + Q(k-1)) \tag{4}$$

$$K(k) = (AP(k-1)A^T + Q(k-1))H^T \left[H(AP(k-1)A^T + Q(k-1))H^T + R(k)\right]^{-1} \tag{5}$$

wherein

$$P(k) = \mathrm{cov}(\hat{x}(k)) \tag{6}$$

$$Q(k) = \mathrm{cov}(w(k)) \tag{7}$$

$$R(k) = \mathrm{cov}(v(k)) \tag{8}$$

Initial conditions on x, P, Q, and R must of course be specified. Quite often P(0) is set to Q(1), Q(1) and R(1) are known, and x(0) may be set to (0).

In the adaptive systems similar to that of Groutage [10], N_s most recent samples of the state error sequence

$$f(k) = \hat{x}(k) - A\hat{x}(k-1) \tag{9}$$

are used to estimate an unknown constant forcing function or bias. That is, let $\hat{u}(k-1)$ be an adaptive state bias estimator,

$$\hat{u}(k-1) = \frac{1}{N_s} \sum_{j=1}^{N_s} f(k-j). \tag{10}$$

This bias is then included in the state estimator:

$$\hat{x}(k) = A\hat{x}(k-1) + \hat{u}(k-1) + K(k)(z(k) - H(A\hat{x}(k-1) + \hat{u}(k-1)) - \hat{r}(k)). \tag{11}$$

Similarly in (11), $\hat{r}(k)$ is an adaptive measurement bias estimator based on the N_Z most recent errors in the predicted measurements:

$$\hat{r}(k) = \frac{1}{N_Z} \sum_{j=1}^{N_Z} [z(k-j+1) - H(A\hat{x}(k-j) + \hat{u}(k-j))] \tag{12}$$

Myers and Tapley [11] show that for H linear the bias measures (10) and (12) are unbiased. Groutage [10] also shows that for linear H that the covariance estimator

$$\hat{Q}(k) = \frac{1}{N_s - 1} \sum_{j=1}^{N_s} [f'(k-j+1) - \hat{u}'(k))(f'(k-j+1) - \hat{u}'(k))^T] - \frac{1}{N_s} \sum_{j=1}^{N_s} (A\hat{P}(k-j)A^T - \hat{P}(k-j+1)) \tag{13}$$

where

$$f'(k) = \hat{x}(k) - A\hat{x}(k-1) - \hat{u}(k-1) \tag{14}$$

and

$$\hat{u}'(k) = \frac{1}{N_s} \sum_{j=1}^{N_s} f'(k-j+1) \tag{15}$$

is unbiased. Similarly the measurement covariance estimate,

$$\hat{R}(k) = \frac{1}{N_Z-1} \sum_{j=1}^{N_Z} \{(y(k-j+1) - \hat{r}(k))(y(k-j+1) - \hat{r}(k))^T - \frac{N_Z-1}{N_Z} H\bar{P}(k-j+1)H^T\}, \quad (16)$$

where

$$\bar{P}(k) = A\hat{P}(k-1)A^T + \hat{Q}(k-1) , \quad (17)$$

$\hat{P}$ is an estimate of P using $\hat{Q}$ in (4), and $y(k) = z(k) - H(A\hat{x}(k-1) + \hat{u}(k-1))$ is shown to be unbiased. Also in [11] are recursive formulations for $\hat{u}$, $\hat{Q}$, $\hat{r}$, and $\hat{R}$.

Although $\hat{r}$ may be estimated using (12), and it is even necessary in estimating R, we show in Appendix A that both $\hat{u}$ and $\hat{r}$ cannot be simultaneously estimated as indicated.

Groutage has proposed modifying the gain K(k) in (5) in such a way that when $\hat{R} \to 0$, the modified gain is the same as (5), but as $\hat{R} \to \infty$, the modified gain changes exponentially rather than inversely toward zero. His example, using noise plus a square pulse of acceleration driving a double integrator and knowing only measurements of position, shows good improvement over both the conventional and previous adaptive techniques. Another Groutage innovation uses the biweight [18] to smooth resulting estimates of states and acceleration input as computed from the derivative of the velocity estimate.

The foregoing ideas have encouraged further nonlinear methods for dealing with unknown deterministic inputs and non-Gaussian noise. We describe these in the following section.

III. ROBUST METHODS FOR NON-GAUSSIAN NOISES AND FORCING FUNCTIONS

The processings discussed in section II are the basis for many tracking schemes. The theoretical developments have evolved from assumptions of linear or linearized state and measurement equations with known deterministic plus known Gaussian inputs to considerations of and allowance for nonlinear equations, unknown non-Gaussian noises and unknown inputs. Exactly at which steps or to which data the robust, adaptive techniques should be applied has not been totally answered, and indeed maybe no one process will be found that answers all problems. Yet it is clear that in the face of all the unknowns, adaptive methods and robust data handlers offer improvements.

The particular problem addressed here is the state system of Equation 1 with unknown step input u(k) and measurements as in Equation 2. The state estimator and its covariance matrix are as in Equations 11 and 4 respectively, but estimates of $\hat{u}$, $\hat{r}$, $\hat{Q}$, and $\hat{R}$ and to be obtained as follows. Thus we have the estimate:

$$\hat{x}(k) = A\hat{x}(k-1) + \hat{u}(k-1) + \hat{K}(k)c(k)[z(k) - H(A\hat{x}(k-1) + \hat{u}(k-1)) - \hat{r}(k)] \tag{18}$$

the measurement:

$$z(k) = H(x(k)) + r(k) + v(k) \tag{19}$$

the error covariance estimate:

$$\hat{P}(k) = (I-K(k)H)(A\hat{P}(k-1)A^T + \hat{Q}(k-1)) \tag{20}$$

the robust process covariance estimate:

$$\hat{Q}(k) = (\hat{q}_{ij}) \tag{21}$$

the robust measurement error covariance estimate:

$$\hat{R}(k) = (\hat{r}_{ij}) \tag{22}$$

the robust forcing function estimate (med = median) based on the N_s most recent estimation errors:

$$\hat{u}(k) = \text{med}\ \{\hat{x}(k) - A\hat{x}(k-1)\}_{N_s} \tag{23}$$

the robust measurement-bias estimate based on the N_z most recent measurement-prediction errors:

$$\hat{r}(k) = \text{med}\ \{z(k) - H(A\hat{x}(k-1) + \hat{u}(k-1))\}_{N_z} \tag{24}$$

and non-optimal Kalman gain:

$$K(k) = (A\hat{P}(k-1)A^T + \hat{Q}(k-1))H^T(k)[H(A\hat{P}(k-1)A^T + \hat{Q}(k-1))H^T + \hat{R}(k)]^{-1} \tag{25}$$

and gain modifier c(k) to be formulated in the following.

The elements $\hat{q}_{im}$ are found by one of the procedures suggested in Mosteller and Tukey [16]. For i = m,

$$\hat{q}_{ii} = \frac{N_s \sum_j{}' (f'(k-j+1) - \text{med}\{f'(j)\})^2 (1-M_{ij}^2)^4}{[\sum_j{}' (1-M_{ij}^2)(1-5M_{ij}^2)][-1 + \sum_j{}' (1-M_{ij}^2)(1-5M_{ij}^2)]} \tag{26}$$

where $(1-M_{ij}^2)^2$ are the biweights for the $j^{\underline{th}}$ sample of the $i^{\underline{th}}$ state variable, and

$$M_{ij} = \frac{f'(k-j+1) - \text{med}\{f'(j)\}}{6\ \text{med}\{|\ f'(m) - \text{med}\{f'(m)\}|\}_{k-N_s+1 \le m \le k}}, \tag{27}$$

and Σ' indicates summation only over those j for which $M_{ij}^2 \le 1$. Thus the biweight weights less those variations of the state estimate which are more distant from the median and gives weight zero those variations which are greater than 6 times the median variation. It can be seen that for small Gaussianly distributed variations around the true median, $\hat{q}_{ij}$ reduces to an unbiased variance estimator.

The covariance estimators $\hat{a}_{im}$ may be found using a number of robust methods. Spearings rank correlation has been shown to be effective, giving

$$\hat{a}_{im} = (\hat{q}_{ii}\hat{q}_{mm})^{\frac{1}{2}}\hat{p}_{im} , \tag{28}$$

where $\hat{p}_{im}$ is the correlation estimate.

The $\hat{r}_{im}$ are found in the same manner as the $\hat{q}_{im}$. For $i = m$,

$$\hat{r}_{ii} = \frac{N_Z \sum_j' (y_i(k-j+1) - \hat{r}_i(k)^2 (1-\eta_{ij}^2)^4}{[\sum_j' (1-\eta_{ij}^2)(1-5\eta_{ij}^2)][-1+\sum_j' (1-\eta_{ij}^2)(1-5\eta_{ij}^2)]} , \tag{29}$$

where

$$\eta_{ij} = \frac{y_i(k-j+1) - \hat{r}_i(k)}{6\ \mathrm{med}\{|y_i(m) - \hat{r}_i(k)|\}\Big|_{k-N_o+1 \le m \le k}} , \tag{30}$$

and

$$y_i(k) = z_i(k) - H(A\hat{x}(k-1) + \hat{u}(k-1)) \tag{31}$$

With the η_{ij} we propose a biweight gain modifier, $c(k)$, a diagonal matrix with diagonal elements

$$c_{ii}(k) = \begin{cases} (1-\eta_{il}^2)^2 , & \eta_{il}^2 \le 1 \\ 0 , & \eta_{il}^2 > 1 \end{cases} \tag{32}$$

Thus for nominal estimated measurement errors $y_i(k)$ with respect to the running error median $\hat{r}_i(k)$, the weight $(1-\eta_{il}^2)^2$ is approximately unity and $c_{ii}(k)$ has no effect on the gain $\hat{K}(k)$. A wild-point measurement z_i, however, causes η_{il}^2 to be large and $c_{ii}(k) = 0$, resulting in the $\hat{x}_i(k)$ becoming $A\hat{x}(k-1) + \hat{u}(k-1)$, simply the prediction.

The formulas (18)-(32) constitute a robust and resistant Kalman filter particularly suited for systems having heavy-tailed and biased non-Gaussian measurement noise and non-deterministic step forcing functions. The estimators $\hat{Q}$ and $\hat{R}$ and gain modifier $c(k)$ will deal with wild points and heavy-tailed or mixed-density noise. The estimate $\hat{u}$ will track steps in unknown

forcing functions; preserving the sharpness of the step yet excluding $(N_s-1)/2$ or fewer consecutive, clustered wild points from affecting the estimation. Example theoretical effects comparing linear and robust estimators near steps are given in Appendix B.

IV. SIMULATION RESULTS OF FIRST-ORDER SYSTEM KALMAN FILTERING

The findings of previous chapters have been integrated into a Kalman filtering problem with a single state. The state and measurement equations are

$$x_{k+1} = Ax_k + u_k + w_k$$

$$z_k = x_k + v_k$$

For all of the simulations we have used (where $N(\mu,\sigma^2) \Rightarrow$ Gaussian)

$$w_k \sim N(0,0.05)$$

$$v_k \sim 0.95N(0,1) + 0.05N(0,51)$$

$$u_k = \begin{cases} 0 & , \quad k \leq 199 \\ 10.0, & k \geq 200 \end{cases}$$

N = window length = 15, $N = N_s = N_z$.

Initial conditions for the three filters tested (conventional, linear adaptive, and robust adaptive) are the same:

$$\hat{x}_0 = 0$$

$$\hat{u}_0 = 0$$

$$\hat{Q} = 0.05$$

$$\hat{R} = 1.00$$

Two values of A are used, 0.1 and 0.9. A = 0.1 gives a wider bandwidth system, following the input shape more closely. The noise-free system output with a step Δu input at $k = 0$ is

$$x_k = \frac{\Delta u}{1 - A} \quad (1 - e^{-k\ell nA})$$

For small process and measurement noises compared to these values, errors in estimates of x and predictions of z after the step will be initially essentially equal to the values of x until the estimate $\hat{u}(k)$ causes the adaptive filter to close the gap. The delay with the robust method is greater, but once the switch is made it may catch up faster than the linear adaptive method (see again Figure B2a,b,d). When the x(k) are more nearly a step (this is true when A is smaller), the advantage of robust as just explained matches the theory and Figure 2 better. For any A, however, the robust method deals with the non-Gaussian noise better.

Another aspect considered in the tests is the use of the factor $c_{11} = (1 - \eta_{11}^2)^2$ per Equations (32) and (30-31) for modifying the Kalman gain when a new measurement is deemed "wild." Another delay is introduced by using c_{11}, but, except near the step, results improve.

One measure of improvement is given by the merit formula

$$M = \sigma_v^2/(\sigma_{\tilde{x}}^2 + \sigma_v^2) \qquad (33)$$

so that as the error variance $\sigma_{\tilde{x}}^2$ approaches zero, M approaches unity. This measure allows comparison of the different process methods.

We rely on the median to estimate the input u_k per Equation (23),

$$\hat{u}_k = \text{med}\{\hat{x}_k - A\hat{x}_{k-1}\}_N .$$

Observe that the residual $z_k - A\hat{x}_{k-1} - \hat{u}_{k-1}$ will be nearly u_k for the first $(N - 1)/2$ samples following a step. This gives a half-window delay (see Figure B2-d) in trying to estimate x_k and trying to estimate

$$\hat{r}(k) = \text{med}\{z_k - A\hat{x}_{k-1} - \hat{u}_{k-1}\} .$$

(Although we do not find $\hat{r}(k)$ to use as a bias on z (we cannot - see Appendix A), it is helpful in estimating $\hat{R}$ and giving a feasible K near the step.)

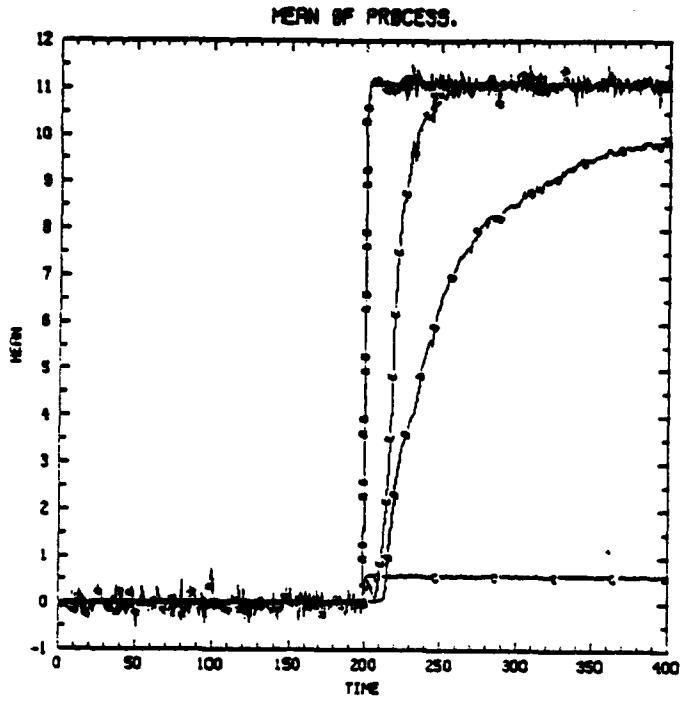
$$R_k = \frac{N \sum_j' (z_{k-j+1} - A\hat{x}_{k-j} - \hat{u}_{k-j} - \hat{r}_k)^2 (1 - \eta_j^2)^4}{[\sum_j' (1 - \eta_j^2)(1 - 5\eta_j^2)][-1 + \sum_j' (1 - \eta_j^2)(1 - 5\eta_j^2)]} \qquad (35)$$

Because of the primed-sums, only those terms which are within the range with $(1 - \eta_j^2) > 0$ are used. This leads to "switching" effects in $\hat{Q}$ and $\hat{R}$. The first switch occurs when the large residuals become the majority; another switch occurs when the state estimates catch up again to the steady-state state and the small residuals are again the majority. Some of these effects can be minimized by refiltering the past (N-1)/2 samples immediately upon step detection, but that feature has not yet been implemented in our work.

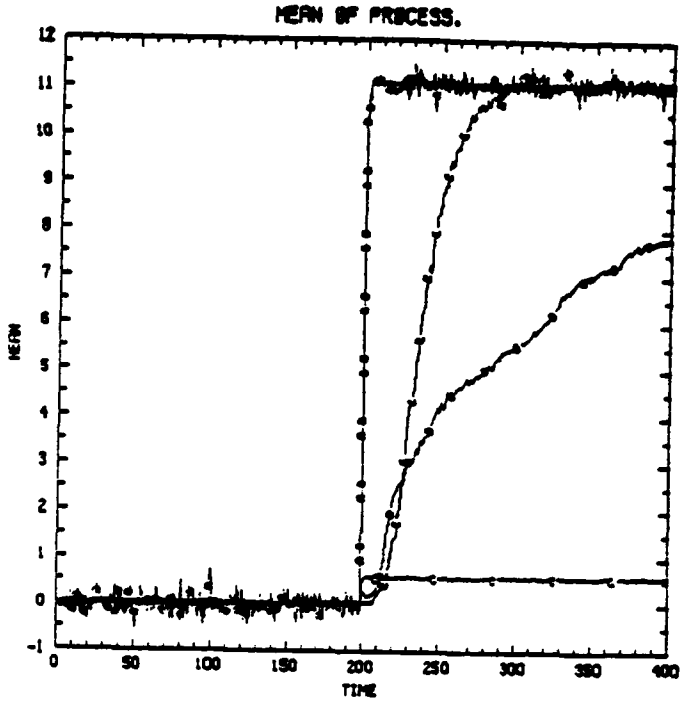
Similar effects occur in the conventional linear adaptive formulas for $\hat{r}$, $\hat{u}$, $\hat{Q}$ and $\hat{R}$, but because of the linearity more gradual errors and also corrections in these estimates occur and peak at different times. Overall system instabilities might be analyzed through the use of the flow-graph in Figure A1.

Although r_k, the "unknown" bias on z, is known here to be zero and is not used in calculating $\hat{x}_k$, $\hat{r}_k$ is used in calculating $\hat{R}$. This prohibits $\hat{R}$ from getting so large that the Kalman filter relies almost entirely on prediction. Again, as the estimate catches up to the true state after the step, z_k and $A\hat{x}_{k-1} + \hat{u}_{k-1}$ are on the same order, $\hat{r}_k$ is nearly zero again and $\hat{R}$ begins to approach its true value.

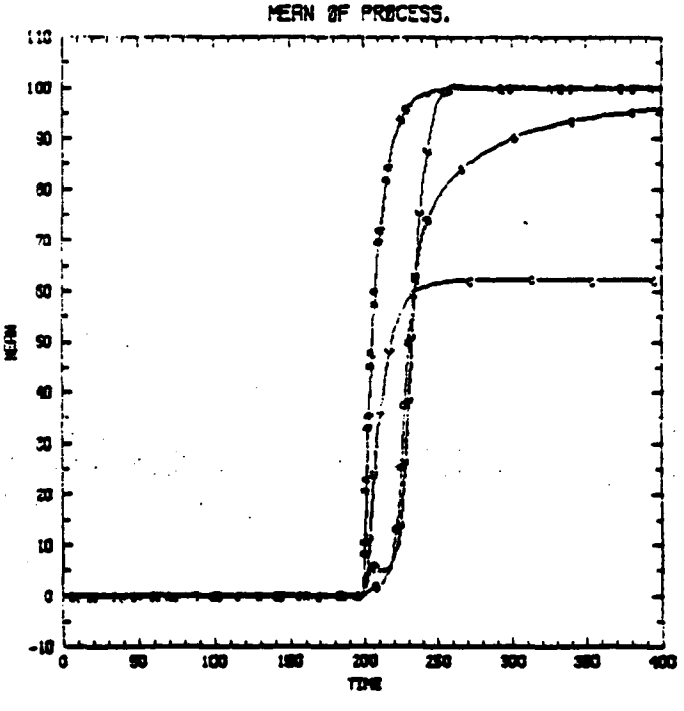
Figure 1 shows the average process outputs; curves A = measurement, B = state, C = conventional filter, D = linear adaptive filter and E = robust adaptive filter. In all cases except for A = .9, $c_{11} = (1 - \eta^2)^2$, the robust estimate recovers more quickly in response to the estimate. In Figure 1d, the combination of the output state's slow rise-time (A = .9) and delay due to $c_{11} = (1 - \eta^2)^2$ causes the robust filter to have considerable delay; however, upon recovery its merit is still superior to the conventional adaptive system because of its ability to



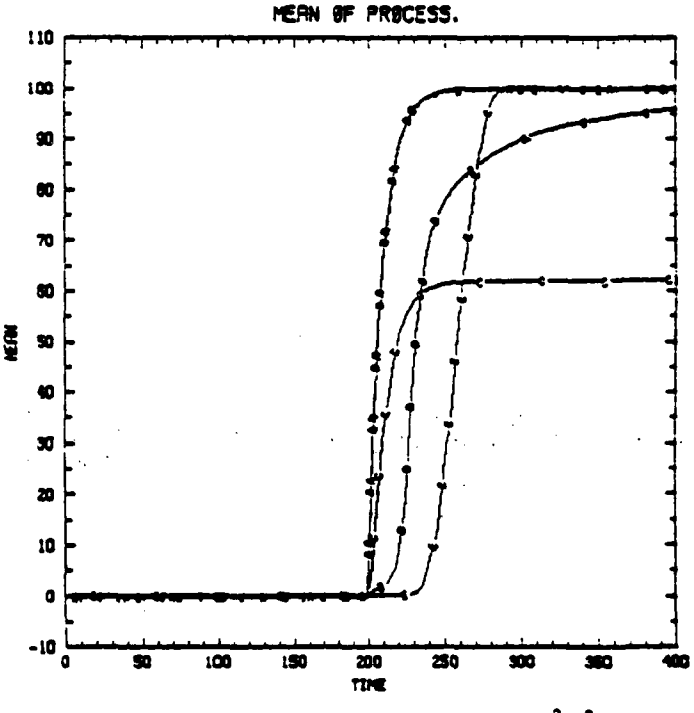
a) A = 0.1, $c_{11} = 1$



b) A = 0.1, $c_{11} = (1 - \eta^2)^2$



c) A = 0.9, $c_{11} = 1$



d) A = 0.9, $c_{11} = (1 - \eta^2)^2$

Figure 1. Average measurement (A), true state (B), and filter estimates (C), linear adaptive (D), and robust adaptive (E).

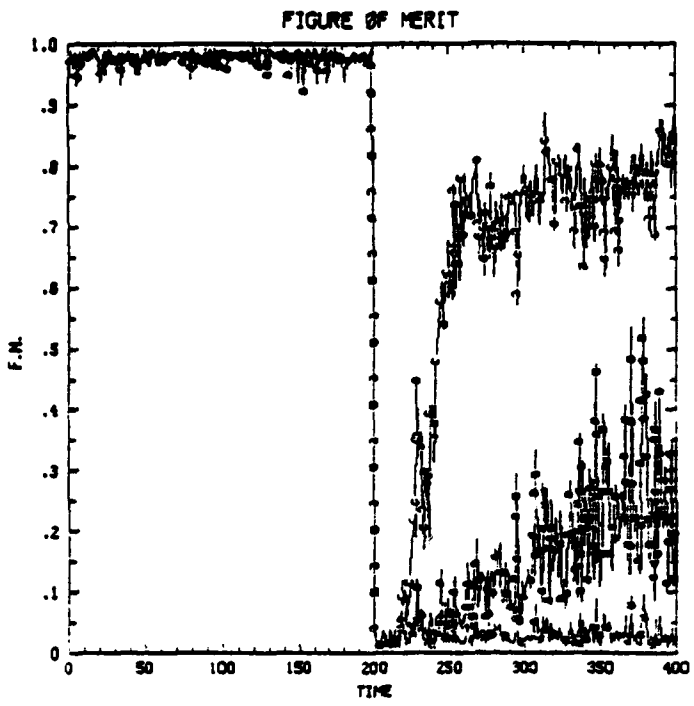
handle the non-Gaussian measurement noise. Figure of merit curves for the four cases of Figure 1 are shown in Figure 2.

In Figure 2 (and all subsequent Figures), curves A, B, C are respectively labels of conventional, linear adaptive, and robust adaptive. In all cases the robust method returns more quickly to a high merit rating. For both the fast and slow systems (A = 0.1 and A = 0.9) the gain biweight factor c_{11} gives an improved transient figure of merit compared to the unmodified Kalman gain K. However some delay is the cost of this improvement. That is, for example in the A = 0.1 system with $c_{11} = 1$, recovery occurs at about k = 230 and robust merit rises to about 0.8 at 400 seconds. But with $c_{11} = (1 - \eta^2)^2$, recovery occurs approximately at k = 270, yet merit quickly rises to about 0.97.

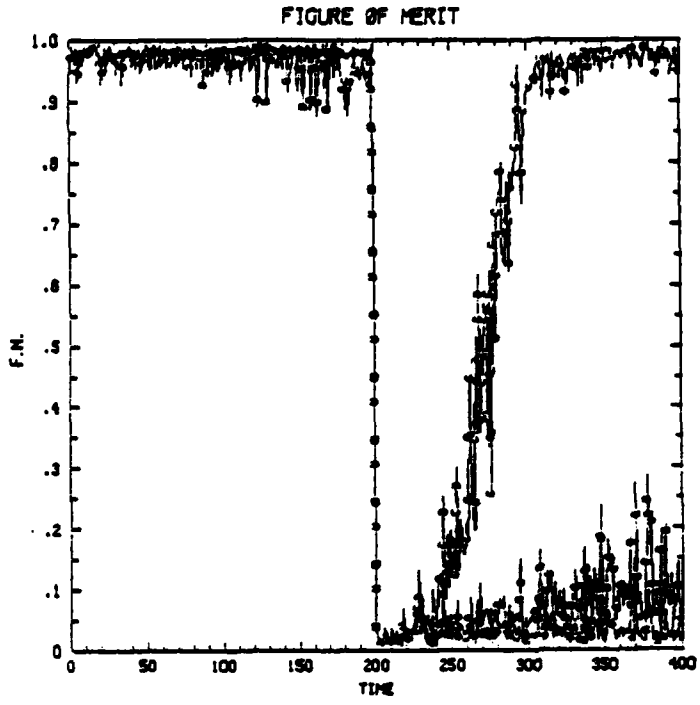
In none of the cases does the average linear adaptive merit figure rise to more than about 0.25 by k = 400.

Figures 3a and 3b are single-run results corresponding to Figures 1b and d respectively (A = 0.1, $c_{11} = (1 - \xi^2)^2$; and A = 0.9, $c_{11} = (1 - \eta^2)^2$), except that the error rather than actual response is shown. Note how in the fast system (Figure 3a) the robust method recovers most quickly, as expected from the ideal-step theory; but in the slow system (Figure 3b) the compounded delays of the robust method are detrimental. Averages of these response types are given in Figures 3c and d. These figures show that the robust average error value is closer to zero from about k = 240 or k = 280, accounting for its higher merit. The non-Gaussian errors evident in Curve B for the single-run in Figure 3a explain this improvement.

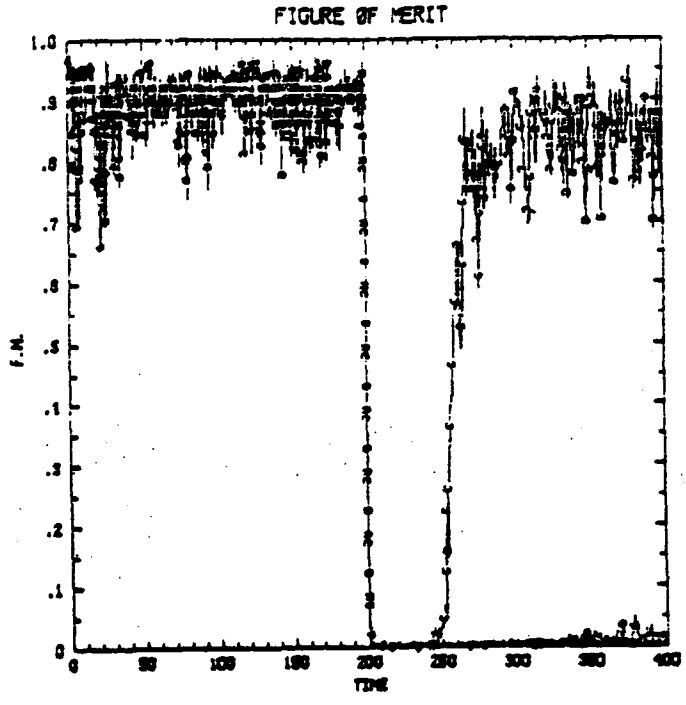
Average estimates of Q and R are shown in Figures 4 and 5 respectively. Considerable variation is noted as parameters vary. However, when the state step is large (A = 0.9), it is clear that $\hat{Q}$ follows true to theory in that during the transition robust $\hat{Q}$ (curve C) deviates less from the uncontaminated



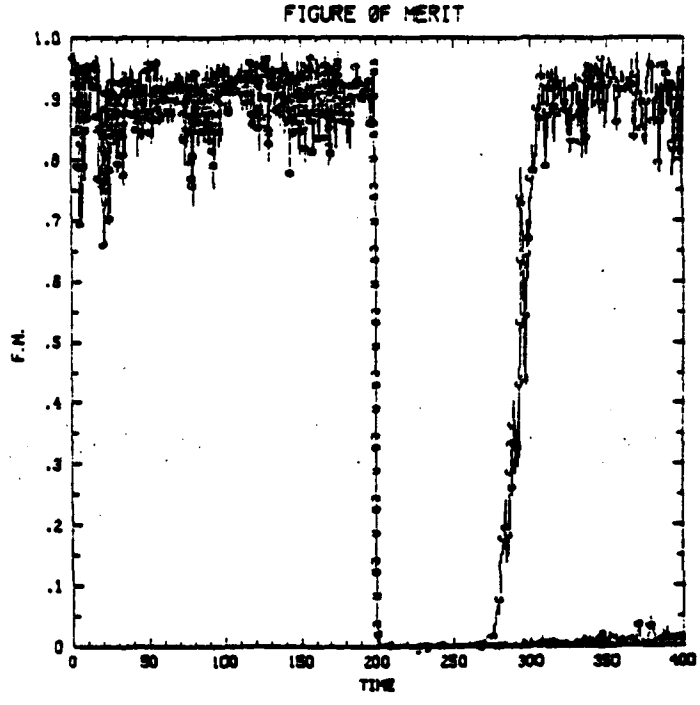
a) A = 0.1, $c_{11} = 1$



b) A = 0.1, $c_{11} = (1 - \eta^2)^2$

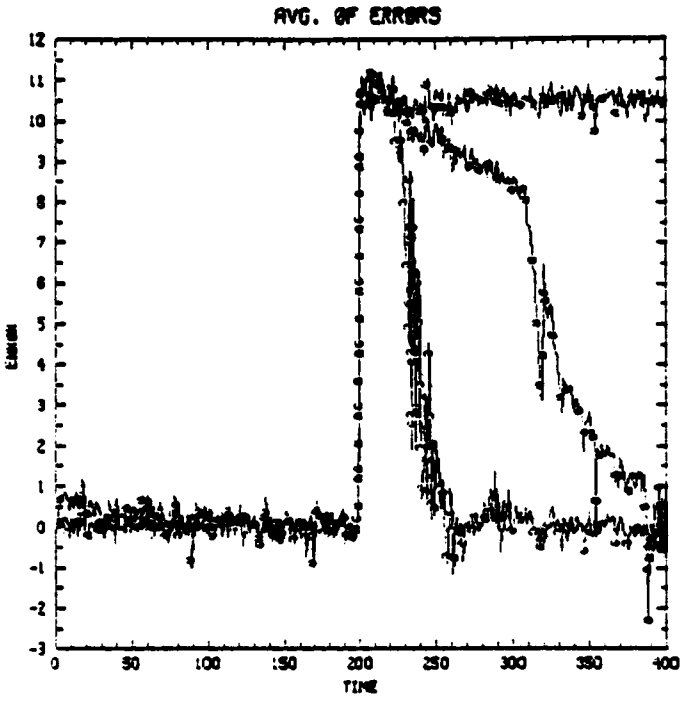


c) A = 0.9, $c_{11} = 1$

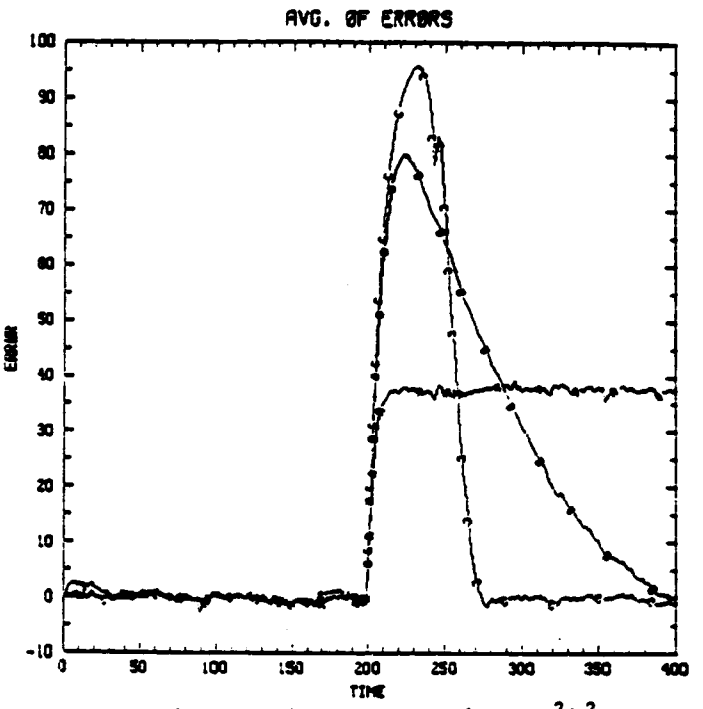


d) A = 0.9, $c_{11} = (1 - \eta^2)^2$

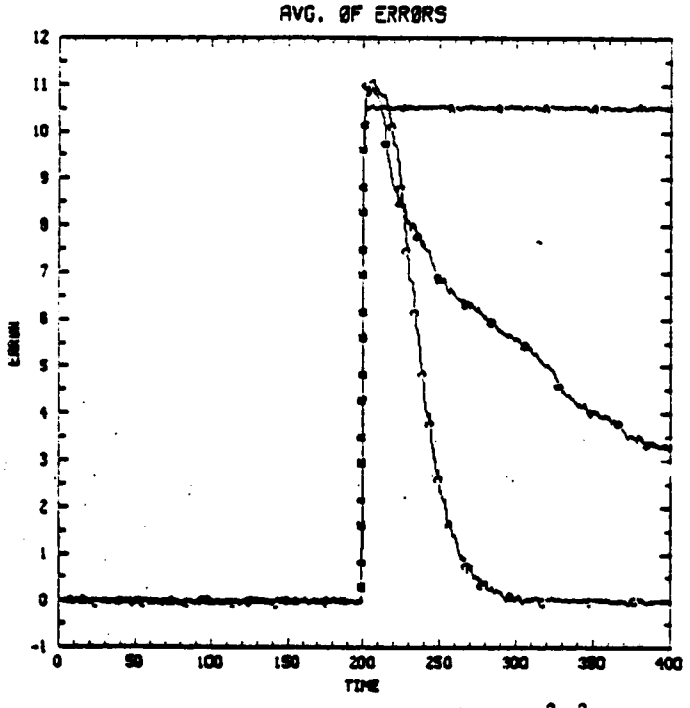
Figure 2. Average merit figures: (A) conventional, (B) linear adaptive, (C) robust adaptive.



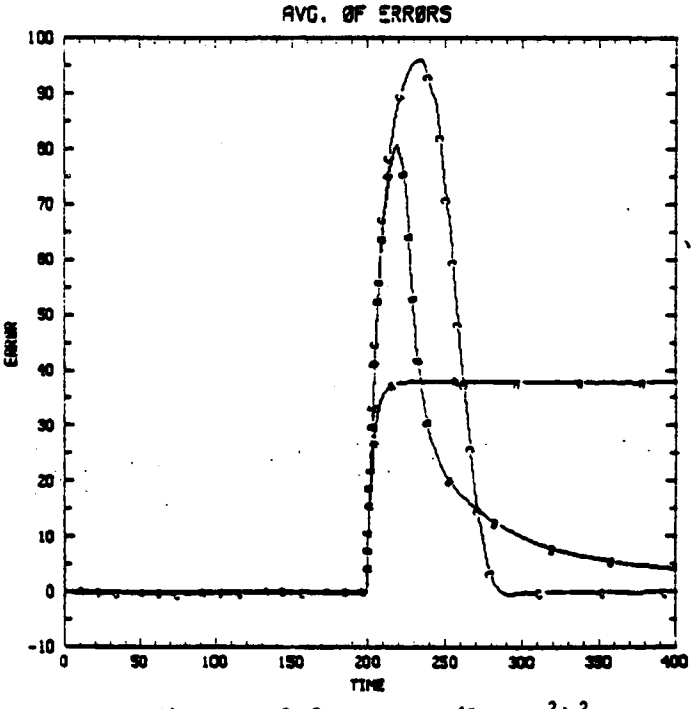
a) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$
single run



b) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$
single run

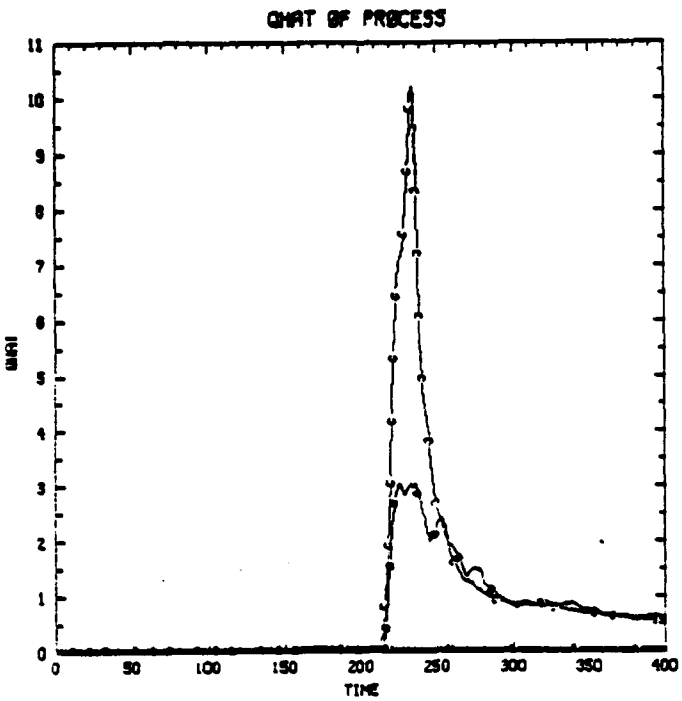


c) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$
average

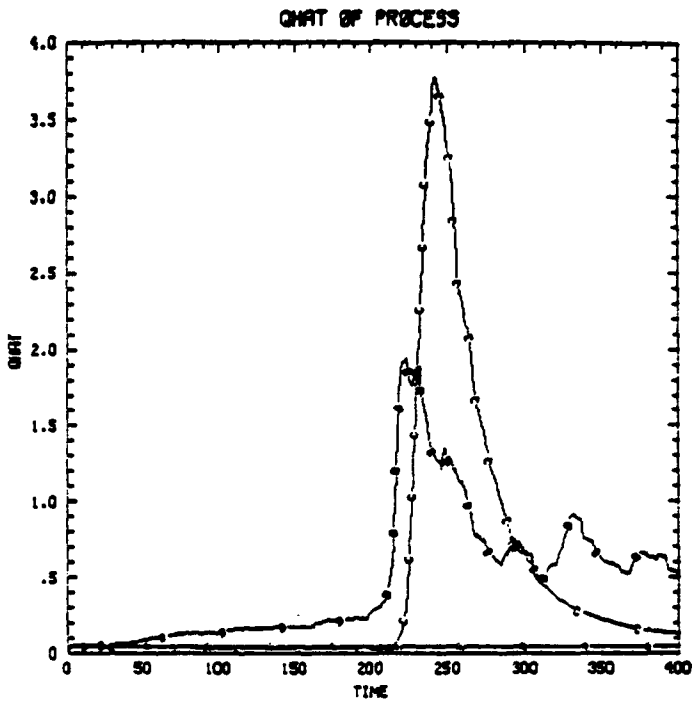


d) $A = 0.9$, $c_{11} = (1 - \eta^2)^2$
average

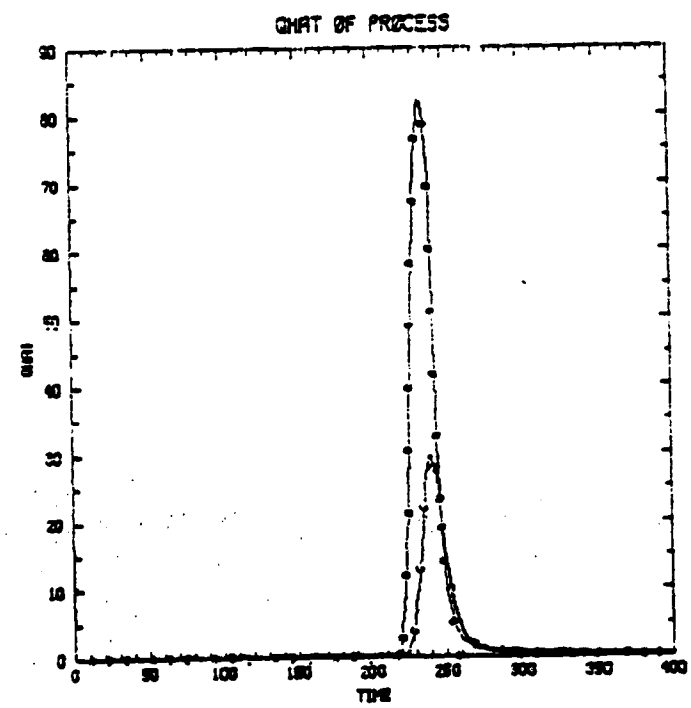
Figure 3. Error plots: (A) conventional, (B) linear adaptive, (C) robust.



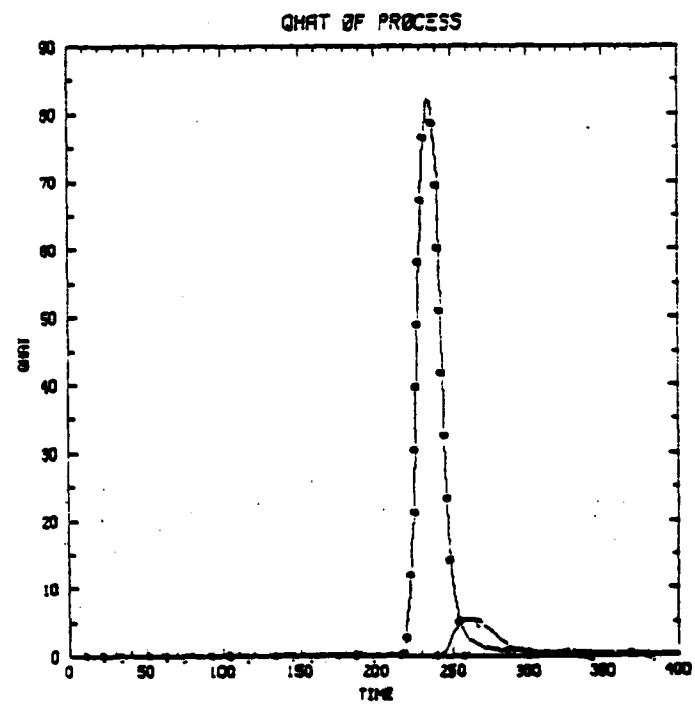
a) $A = 0.1$, $c_{11} = 1$



b) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$

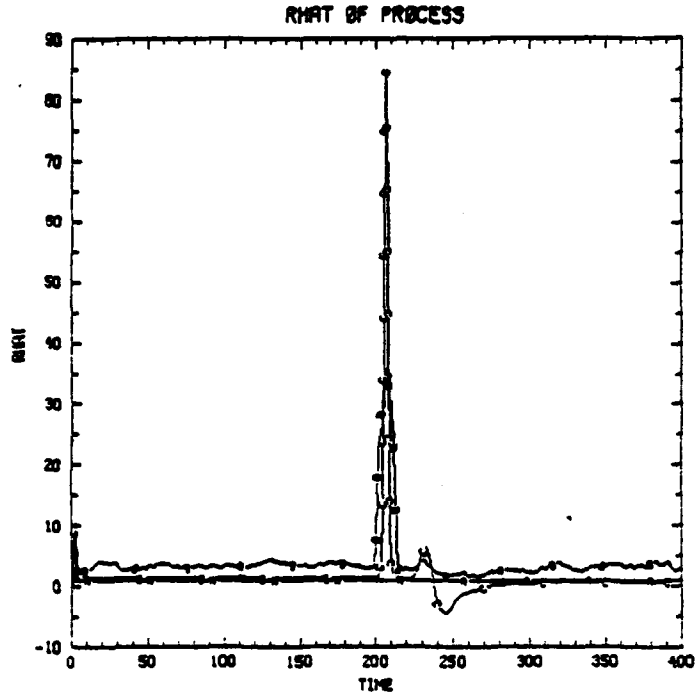


c) $A = 0.9$, $c_{11} = 1$

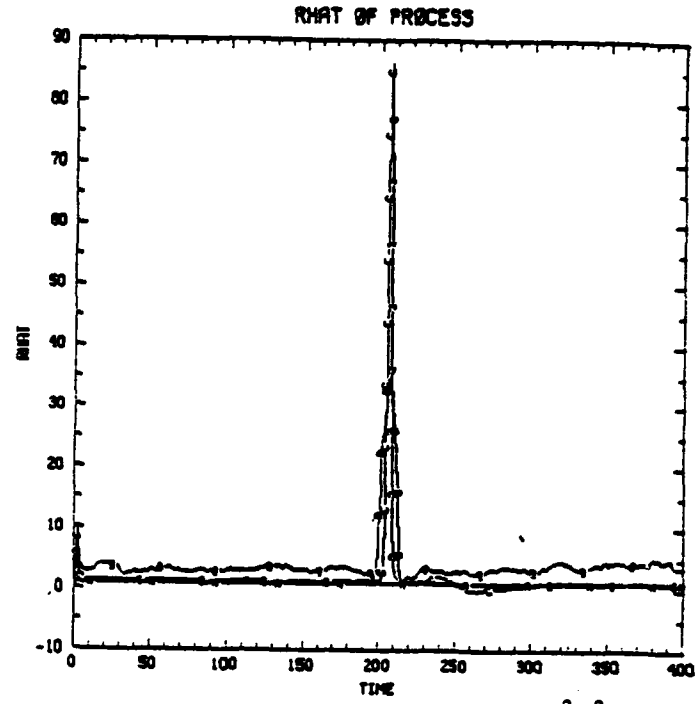


d) $A = 0.9$, $c_{11} = (1 - \eta^2)^2$

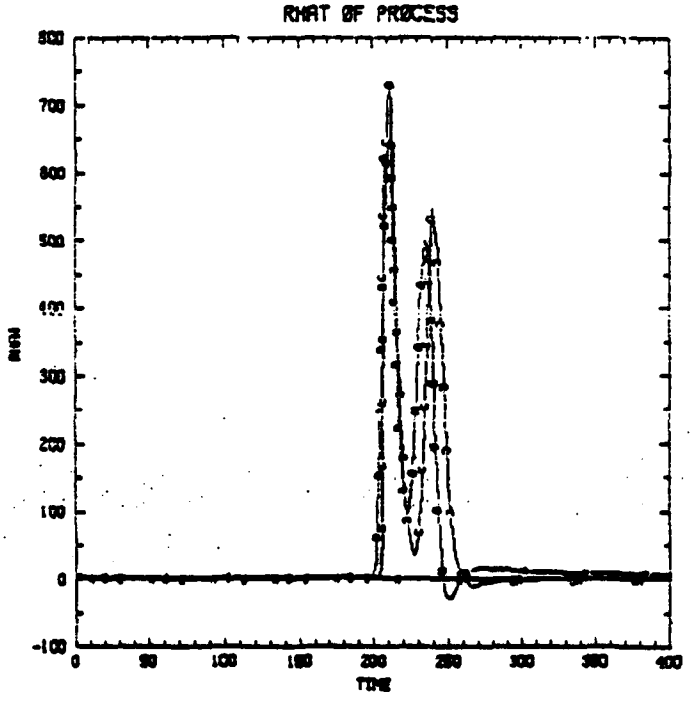
Figure 4. Average $\hat{Q}$: (A) conventional, (B) linear adaptive, (C) robust.



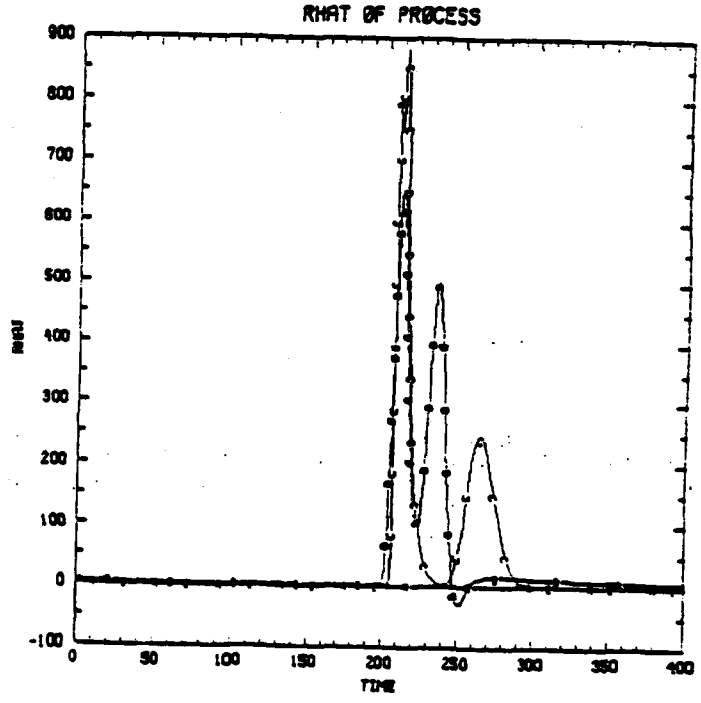
a) $A = 0.1$, $c_{11} = 1$



b) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$



c) $A = 0.9$, $c_{11} = 1$



d) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$

Figure 5. Average $\hat{R}$: (A) conventional, (B) linear adaptive, (C) robust.

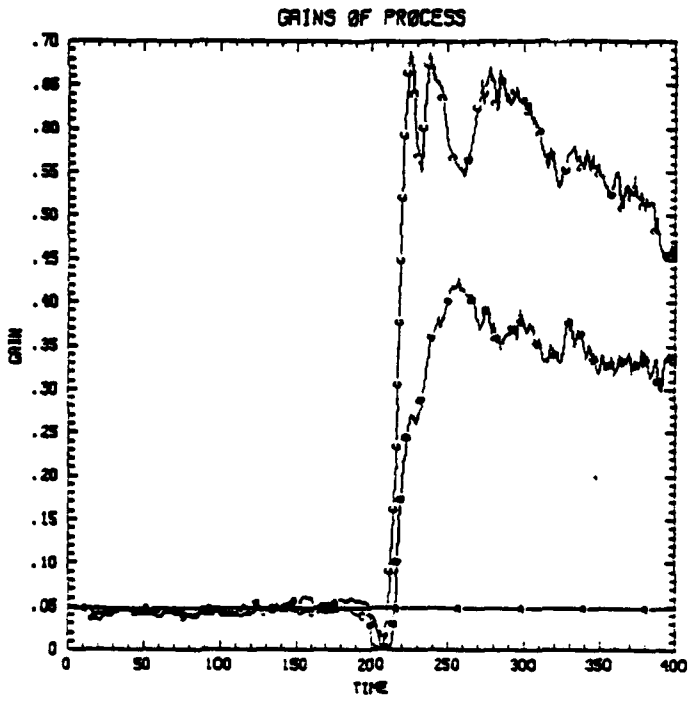
process variance than does the linear $\hat{Q}$ (curve B). Curve A, the conventional non-adaptive Q, of course is always correct. We have not yet tested for the situation where true variance switches values.

The essentially unpredictable effects on average Kalman gains, excluding the c_{11} factor are given in Figure 6. The correct gain for true Gaussian measurement noise and no step is in curves A, and these are constant. Curves B and C should maintain this level before the step and should eventually return to this level.

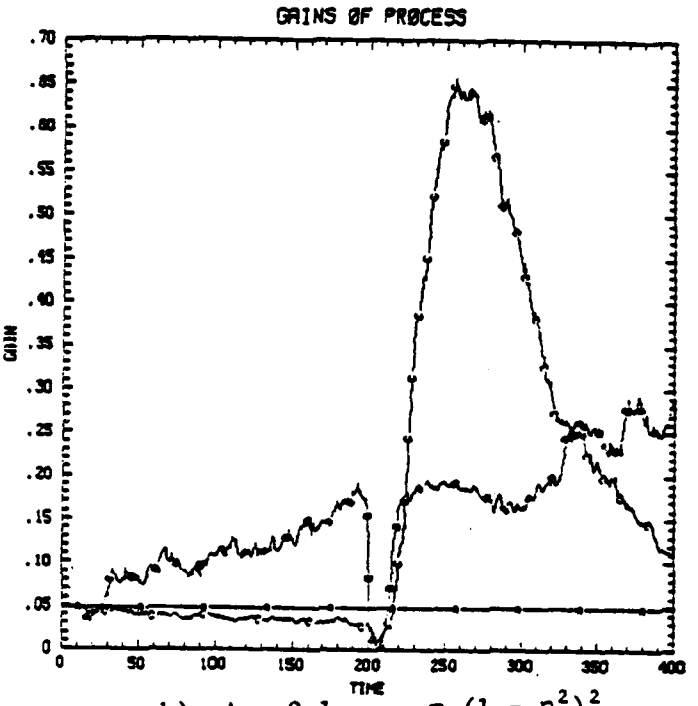
In Figure 7 are the estimates of the forcing function for the four parameter sets. Again, because the state estimation errors in the fast system (Figures 7a-b) are more step-like, the robust method gives a good improvement. Only in Figure 7d is the delay excessive, and even then curve C is better after recovery.

V. CONCLUSIONS

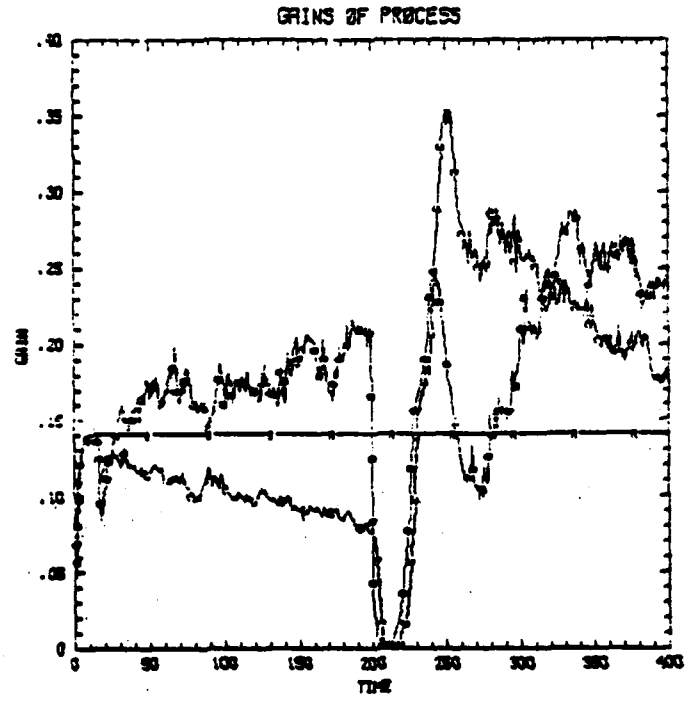
Features of separate previous works have been implemented into one algorithm to estimate states in systems having unknown step inputs which are large with respect to the process noise, and to handle non-Gaussian measurement errors. The running window to observe signal steps in non-Gaussian noise has been robustified by the use of medians and biweights. These statistics have proven useful in making more accurate and responsive the otherwise conventional Kalman filter. A figure of merit which compares state estimation error to measurement error has been introduced. It shows quite well the advantage of the robust scheme over earlier linear adaptive and conventional filterings. From these results we are encouraged to seek more theoretically optimal robust techniques and to experiment with higher order systems and other-than step inputs.



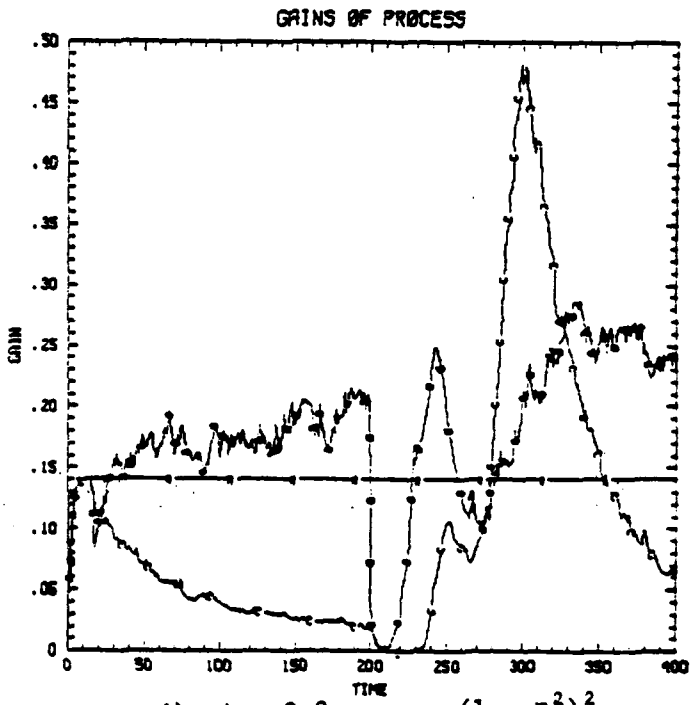
a) A = 0.1, $c_{11} = 1$



b) A = 0.1, $c_{11} = (1 - \eta^2)^2$

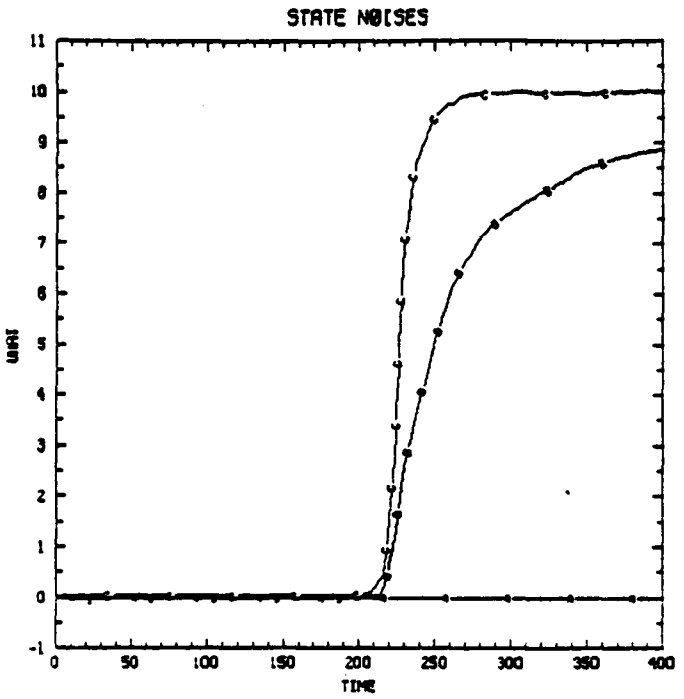


c) A = 0.9, $c_{11} = 1$

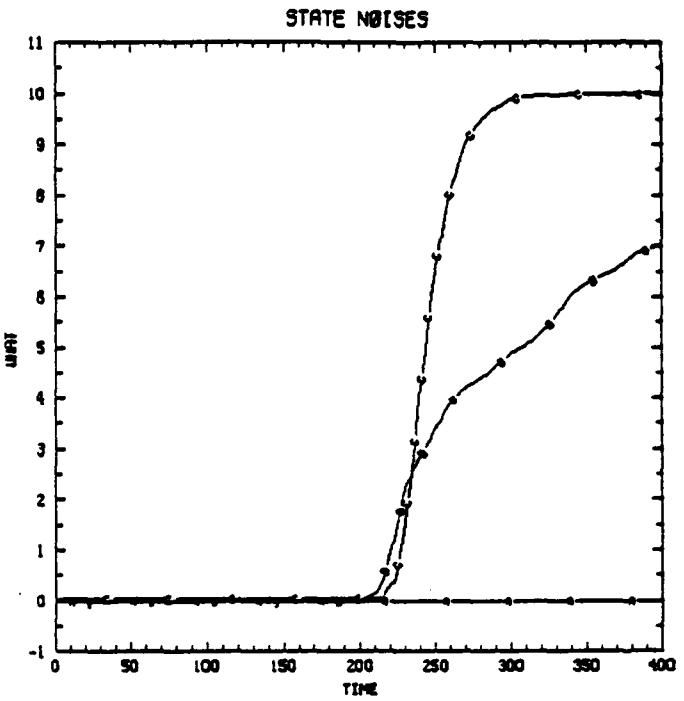


d) A = 0.9, $c_{11} = (1 - \eta^2)^2$

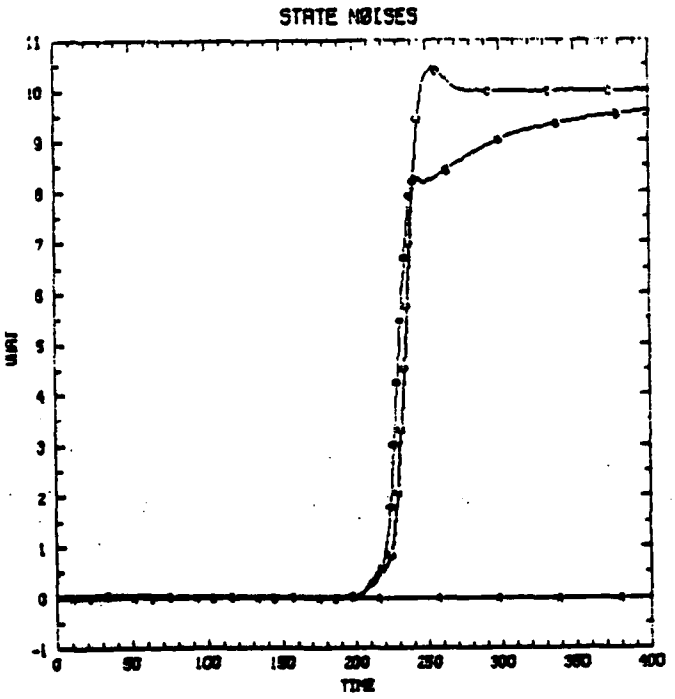
Figure 6. Kalman gains, excluding c_{11}: (A) conventional, B) linear adaptive, (C) robust.



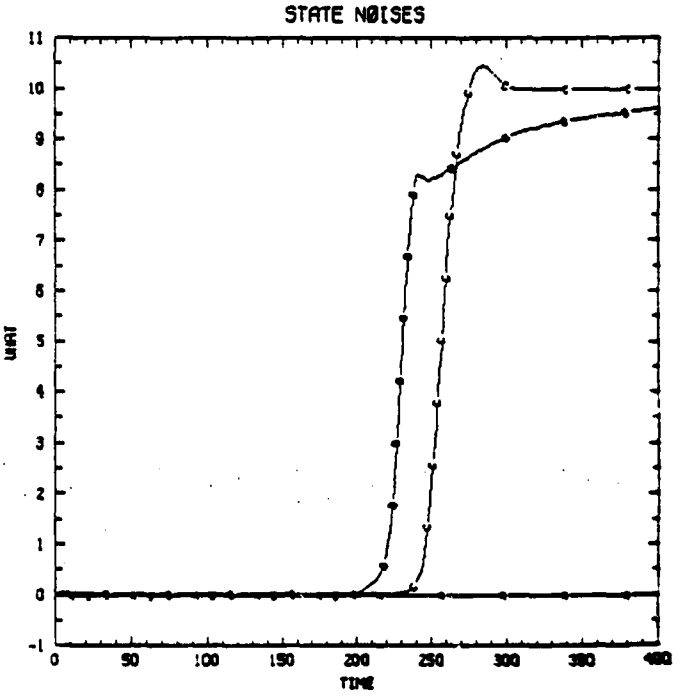
a) $A = 0.1$, $c_{11} = 1$



b) $A = 0.1$, $c_{11} = (1 - \eta^2)^2$



c) $A = 0.9$, $c_{11} = 1$



d) $A = 0.9$, $c_{11} = (1 - \eta^2)^2$

Figure 7. Step input estimators: (A) conventional, (B) linear adaptive and (C) robust.

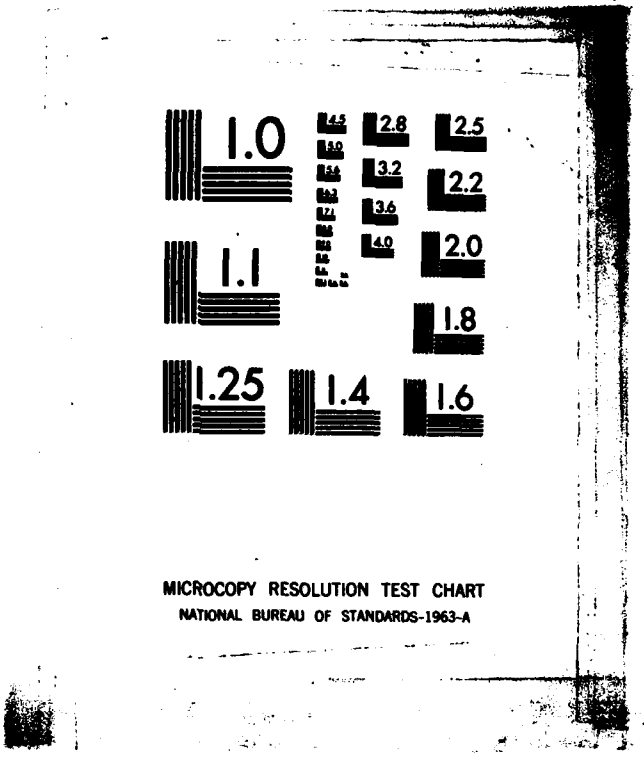
AD-A147 619 ROBUST ADAPTIVE KALMAN TRACKERS FOR SYSTEMS WITH UNKNOWN STEP INPUTS NON-.. (U) WYOMING UNIV LARAMIE R L KIRLIN ET AL. 31 OCT 84 N00014-82-K-0048 2/2

UNCLASSIFIED F/G 12/1 NL

END

FILMED

DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

APPENDIX A

COMMENTS ON THE ADAPTIVE ALGORITHMS

Myers and Tapley [11] proposed a method of estimating an unknown forcing constant $u(k) = u$ and any bias r in the measurement noise. These estimates are (similar to Equations 10, 12-16)

$$\hat{u}(k) = \frac{1}{N_s} \sum_{j=k-N_s+1}^{k} f(j) \qquad \text{A-1}$$

and

$$\hat{r}(k) = \frac{1}{N_z} \sum_{j=k-N_z+1}^{k} y(j) \qquad \text{A-2}$$

where

$$f(k) = \hat{x}(k) - A\hat{x}(k-1) \qquad \text{A-3}$$

and

$$y(k) = z(k) - H(A\hat{x}(k-1) - \hat{u}(k-1)) \qquad \text{A-4}$$

The covariance estimators were given to be

$$\hat{Q} = \frac{1}{N_s-1} \sum_{j=k-N_s+1}^{k} (f(j) - \hat{u}(k))(f(j) - \hat{u}(k))^T - \frac{1}{N_s} \sum_{j=k-N_s+1}^{k} (A\hat{P}(j-1)A^T - \hat{P}(j)) \qquad \text{A-5}$$

and

$$\hat{R} = \frac{1}{N_z-1} \sum_{j=k-N_z+1}^{k} (y(j) - \hat{r}(k))(y(j) - \hat{r}(k)^T - \frac{1}{N_z} \sum_{j=k-N_z+1}^{k} H(A\hat{P}(j-1)A^T + \hat{Q}(j-1))H^T \qquad \text{A-6}$$

Groutage [10] rederived the Q and R estimators but used $f'(k) = \hat{x}(k) - A\hat{x}(k-1) - \hat{u}(k-1) = \hat{x}(k) - \bar{x}(k)$ in $\hat{Q}$. The analysis in [10] leads to

$$\hat{Q} = \frac{1}{N_s-1} \sum_{j=k-N_s+1}^{k} (f'(j) - \hat{u}'(k))(f'(j) - \hat{u}'(k))^T$$

$$- \frac{1}{N_s} \sum_{j=k-N_s+1}^{k} (A\hat{P}(j-1)A^T - \hat{P}(j)) \quad , \qquad \text{A-7}$$

where $u'(k)$ is the average of the $f'(j)$. Averaging 100 simulations has shown that this estimator leads to accurate results, $\hat{Q}$ going negative only rarely. (Myers and Tapley used $|\hat{Q}|$ to overcome negative estimates. This study chooses instead to use the most recent positive estimate). Note that A-1 must still be used to estimate u, the step input.

The biases of the two sample covariance estimators are the same. Also the difference in terms,

$$(f'(j) - \hat{u}'(k)) - (f(j) + \hat{u}(k))$$

$$= -\hat{u}(j-1) + \frac{1}{N_s} \sum_{m=k-N_s+1}^{k} \hat{u}(m-1) \quad , \qquad \text{A-8}$$

has expected value equal to zero. Thus we expect that the difference in variances of $\hat{Q}$ by A-5 and by A-7 is also small or zero.

Next the instability of the variance estimators demonstrated in [11] and remarked on in other literature is considered here. A flowgraph showing system states, the Kalman filter and estimators for r and u is given in Figure A1. It is clear that the constants appearing as inputs at the summing junction whose output is z(k) cannot be distinguished after their addition. Analysis of the iterations in the scalar case verifies the proportioning of

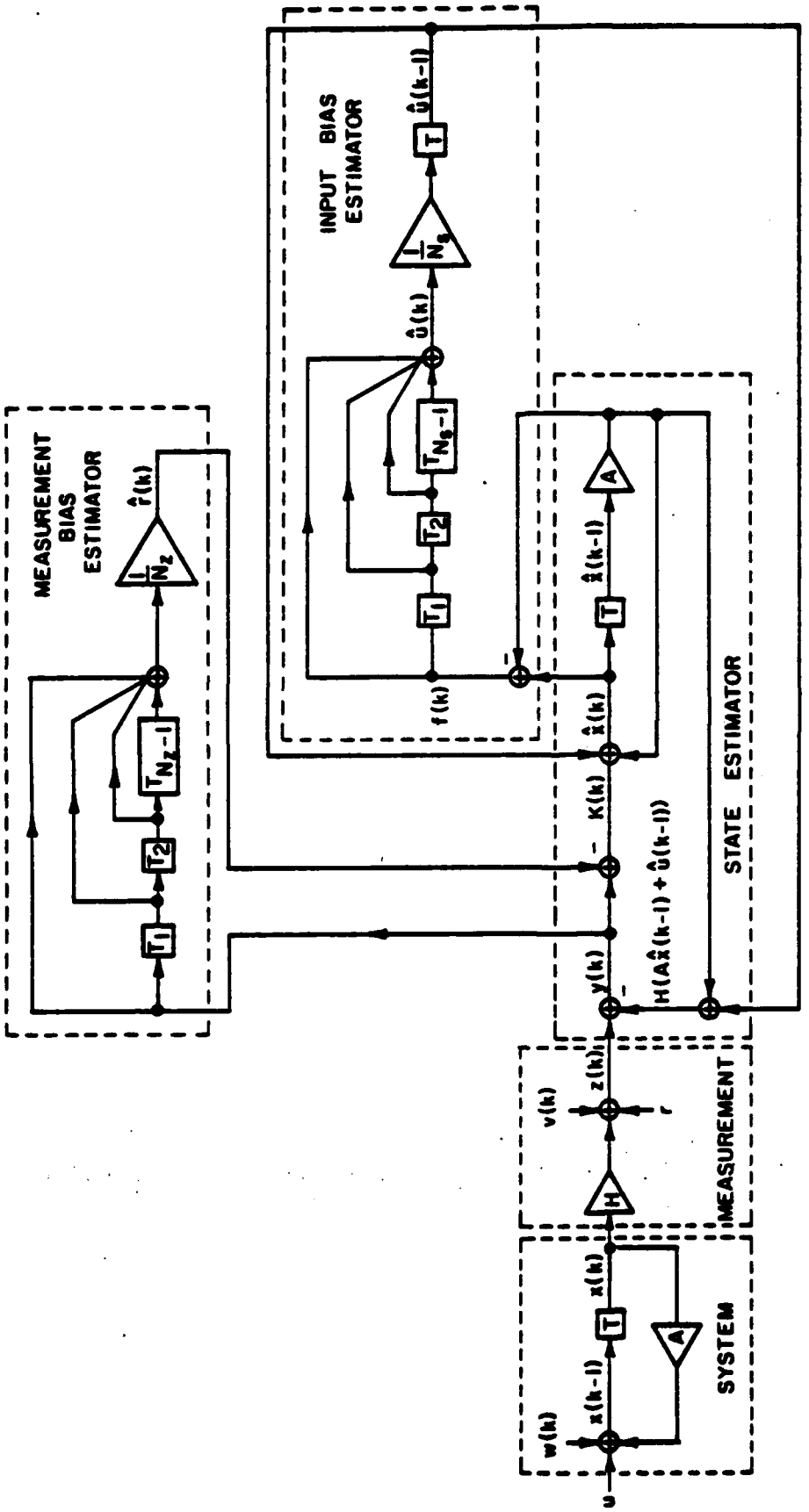


Figure A1. Myers and Tapley Algorithm for estimation of input bias u and measurement bias r.

any one of the biases into the two estimators. This erroneous division of the biases leads to erroneous $\hat{Q}$ and $\hat{R}$ which in turn contaminate the Kalman gain and P matrices. A solution to this difficulty has not yet been proposed other than to know a priori one of the biases - particularly r, as we are particularly interested in unknown step inputs.

APPENDIX B

ERROR ANALYSIS NEAR STEPS

The adaptive estimators for $\hat{u}$, $\hat{r}$, $\hat{Q}$ and $\hat{R}$ formulated in Section III give rise to certain errors. Those of greatest concern here are the transient errors in tracking a step input and the mean square errors in estimating process and measurement noise variances. Because both formulas and numerical data are available in the literature for Gaussian noise, our examples will be taken therefrom. Results will be indicative of those for the non-Gaussian case, but will also clarify the non-optimality of the robust procedures when errors are indeed Gaussian.

First consider a running N-point sample median $\tilde{w}$ on data w_i (Here sequence is denoted by subscripting). Justusson states (in [16]) that the approximate variance if the w_i are independent and identically distributed (i i d) normal N(m, 0) then

$$\text{var}[\,\tilde{w}\,] \simeq \frac{\sigma^2}{N + \pi/2-1} \cdot {}^{\pi}/_{2}\,, \qquad \text{B-1}$$

which is 57% larger than that of the sample mean, $\bar{w}$. However if the noise is distributed with density (Laplacian)

$$f(w) = \frac{\sqrt{2}}{\sigma}\; e^{-\sqrt{2}\;|w-m|/\sigma} \qquad \text{B-2}$$

then

$$\text{var}[\,\tilde{w}\,] \simeq \frac{\sigma^2/2}{N-1/2}\,, \qquad \text{B-3}$$

which is 50% smaller than the variance σ^2/N of the mean. Also for this density, the median is that value a which minimizes

$$\min_{a} \sum_{i=1}^{N} |w_i - a|\,. \qquad \text{B-4}$$

This leads to believing that medians are useful for smoothing noise with heavy-tailed densities. A testament to this belief is found in [20],

wherein it is shown that as noises become more implusive than the Laplacian, the optimal order-statistic filter quickly approaches the median. To begin the error analysis, let a window of length N straddle a step in the data such that the variables $w_1, w_2, \ldots, w_{N-k}$ have density $N(0,\sigma)$ and the variables $w_{N-k+1}, \ldots, w_N$ have density $N(S,\sigma)$, and assume $k < (N+D/2)$. For small S, $E\{\tilde{w}\}=E\{\bar{w}\} = Sk/n$; but for larger S,

$$\tilde{w} \simeq w_{[(N+1)/2,\ N-k]}.$$

where $w_{[(N+1)/2,\ N-k]}$ is the $(N+1)/2^{\underline{th}}$ order statistic of $w_1, w_2, \ldots w_{N-k}$ $(w_i \le w_{i+1})$.

The mean of $\tilde{w}$ is then bounded by [19] $\sigma F^{-1}\left(\frac{(N+1)/2}{N-k+1}\right) \le E\{w_{[(N+1)/2,\ N-k]}\} \le \sigma F^{-1}\left(\frac{N/2}{N-k}\right)$, where F is the cumulative distribution of w.

The analytical expression for the variance of $\tilde{w}$ is available in the literature [19] only in the form

$$\sigma^2_{\tilde{w}} = \int_{-\infty}^{\infty} (\zeta - E\{\tilde{w}\})^2 \, f_{\frac{N+1}{2};\ N-k}(\zeta)\,d\zeta \qquad \text{B-5}$$

where

$$f_{\frac{N+1}{2};N-k}(w) = \frac{(m-1)!\,(N-k-m)!}{(N-k)!}\, F^{m-1}(w)[1-F(w)]^{N-k-m} f(w) \qquad \text{B-6}$$

and $m = (N+1)/2$.

We may thus plot $E\{\tilde{w}\}/S$ and $\sigma_{\tilde{w}}$ vs i, for various S, σ and N, where

$$w_i \sim \begin{cases} N(0,\ \sigma)\ , & i \le 0 \\ N(S,\ \sigma)\ , & i \ge 1\ . \end{cases} \qquad \text{B-7}$$

An example $E\{\tilde{w}\}$ and $\bar{w}$ vs i for σ=1, S=5, N=3 is given by Justusson (in [16]). As the expected error in tracking the step with a running median is symmetrical about i = N/2 (considering N odd, N/2 is between $i = \frac{N-1}{2}$ and $\frac{N+1}{2}$), we need only plot the expected error, $E\{\tilde{w}(j)\}$ where $1 \le j \le \frac{N-1}{2}$.

This is shown in Figure B1 for several values of N, assuming $S \geq 5\sigma$, using data from Teichroew [22]. Shown in Figure B2 are plots of $\tilde{w}$ and $\bar{w}$ for a no-noise situation, and $E\{\hat{Q}\}$ and $E\{\hat{\sigma}_w^2\}$ for S=5, σ=1, N=5. The results show that the median does a very good job of tracking the step in comparison to the mean. This feature offers great advantage over conventional adaptive and non-adaptive methods in Kalman filtering.

In Figure B2 two advantages of the robust technique are evident. First, the step of med$\{w_i\}$ is a delayed but otherwise exact replica of the signal; the overall sum squared error is greater by a factor of 3N/(2N-1), as will be shown. Secondly the running linear standard deviation measure $\hat{\sigma}_w$ has an approximate peak value $S/\sqrt{2}$, whereas the peak robust $\hat{Q}^{1/2}$ has approximate value 2σ (for N=5 and $S \geq 5\sigma$). These values are found as follows.

The sum-squared error of the running sample average in the noise free case is

$$SS_a = \sum_{i=1}^{N-1} (\bar{w}(i)-S)^2$$

$$= S^2(N-1)(2N-1)/(6N) \qquad \text{B-8}$$

The sum-squared error of the running sample median in the noise-free case is

$$SS_m = \sum_{i=1}^{N-1} (\tilde{w}(i)-S)^2$$

$$= (N-1)S^2/2 \qquad \text{B-9}$$

The degradation factor is given by

$$SS_m/SS_a = 3N/(2N-1), \qquad \text{B-10}$$

approximately a factor of 3/2.

The mean of the running linear variance estimator, $\hat{\sigma}_\varepsilon^2$, in the transition region $1 \leq i \leq N-1$ for the low noise case is essentially the mean

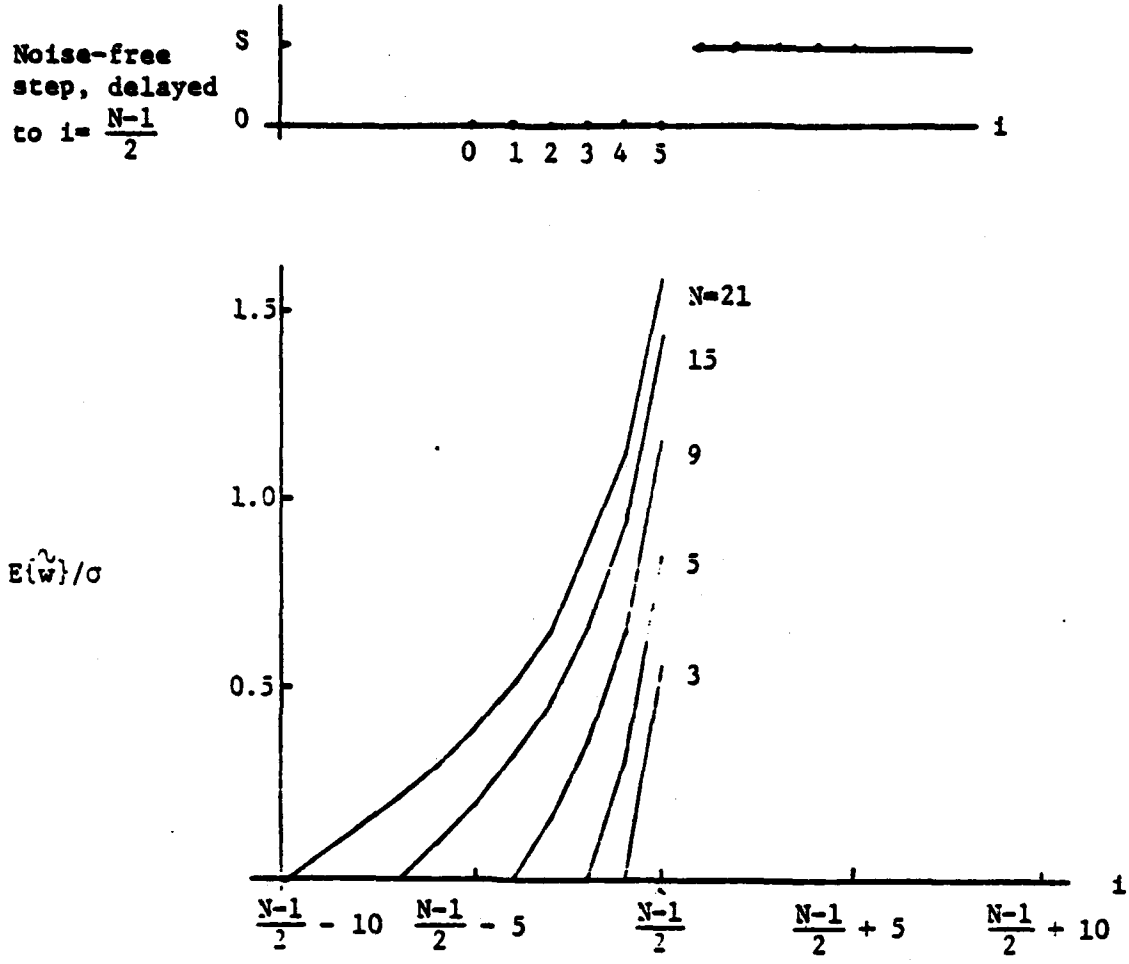


Figure B1. Normalized expected error from the true signal step (size S) for a size N sample set, $S/\sigma \geq 5$. Error is symmetric around i = N/2.

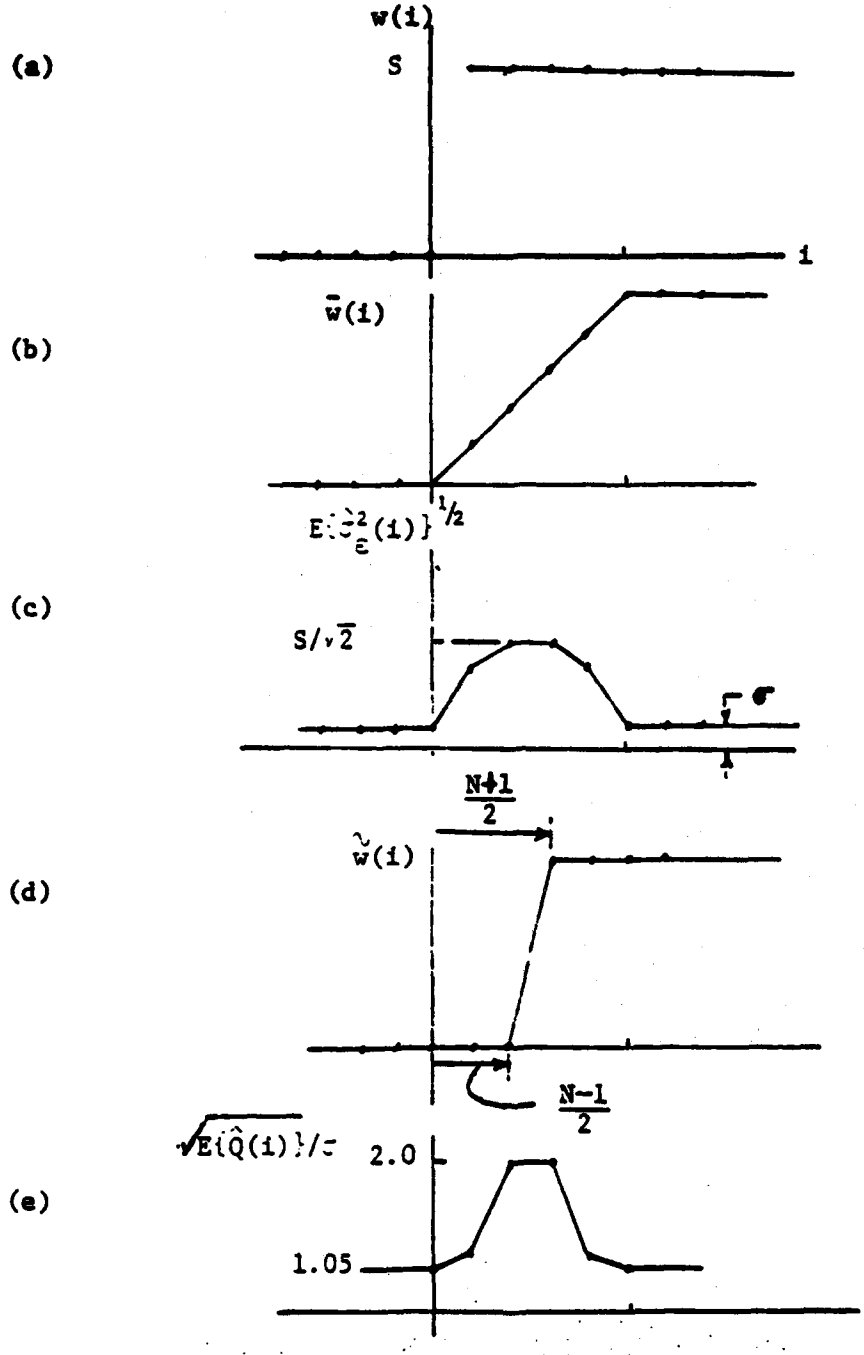


Figure B2. Sequence waveforms of running location and variances estimators; (a) noise-free step in data, (b) 5-point running noise-free average, (c) 5-point running nearly noise-free ($S \geq 5\sigma$) mean variance estimate, (d) 5-point running noise-free median, (e) 5-point running robust $\hat{Q}$, $S \geq 5\sigma$.

square of the error $S - \bar{w}(j)$ over $i-N+1 \le j \le i$. Considering no error for $j \le 0$ or $j \ge N$,

$$E\{\hat{\sigma}_\varepsilon^2(i)\} \approx \frac{1}{N-1} \sum_{j=1}^{i} (S - jS/N)^2, \qquad 1 \le i \le N-1$$

$$= (S^2/N^2)i(N^2 - N(i+1) + (i+1)(2i+1)/6) \qquad \text{B-11}$$

The peak value in B-11 occurs at $i = (N \pm 1)/2$ and it is $S^2/2$. Thus the peak standard deviation estimate is $S/\sqrt{2}$.

The mean of the robust measure of variance, $\hat{Q}$, is plotted in Figure B3, using data from Teichroew [20], under the assumption that S/σ is large. This approximation is derived as follows. In this scalar case $\hat{Q}$ is $\hat{q}_{ii}$ of Equation 26, with $\tilde{w}(i)$ representing the $\hat{u}_i(k)$, k=i, and w(j) representing $\hat{x}_i(k-j+1) - A\hat{x}_i(k-j)$. Thus at time i, and assuming b_{ii} negligible,

$$E\{\hat{Q}(i)\} = E\left\{ \frac{N \sum_j' (w(j) - \tilde{w}(i))^2 (1-\mu_j^2)^4}{[\sum_j' (1-\mu_j^2)(1-5\mu_j^2)][-1 + \sum_j' (1-\mu_j^2)(1-5\mu_j^2)]} \right\} \qquad \text{B-12}$$

where $\sum_j'$ only includes the N most recent data for which $\mu_j^2 \le 1$ and

$$\mu_j = \frac{w(j) - \tilde{w}(i)}{6\ \text{med}\{|w(m) - \tilde{w}(i)|\}\big|_{j-N+1 \le m \le i}} \qquad \text{B-13}$$

In the noise free case, $\text{med}\{|w(m)-\tilde{w}(i)|\}$ is exactly equal to zero because the majority of the w(m) in the window are exactly equal to $\tilde{w}(i)$. Now there are i values nearly equal to S and N-i values nearly equal to 0 within the window, and let $i \le (N-1)/2$. Then $\text{med}\{|w(m)-\tilde{w}(i)|\}$ is very small w.r.t. to S and $w(j) - \tilde{w}(i)$ in the numerator of B-13 will have magnitudes smaller than $6\ \text{med}\{|w(m)-\tilde{w}(i)|\}$ only for N-i of the w(j) samples. Thus $E\{\hat{Q}(i)\}$ reduces for large S/σ

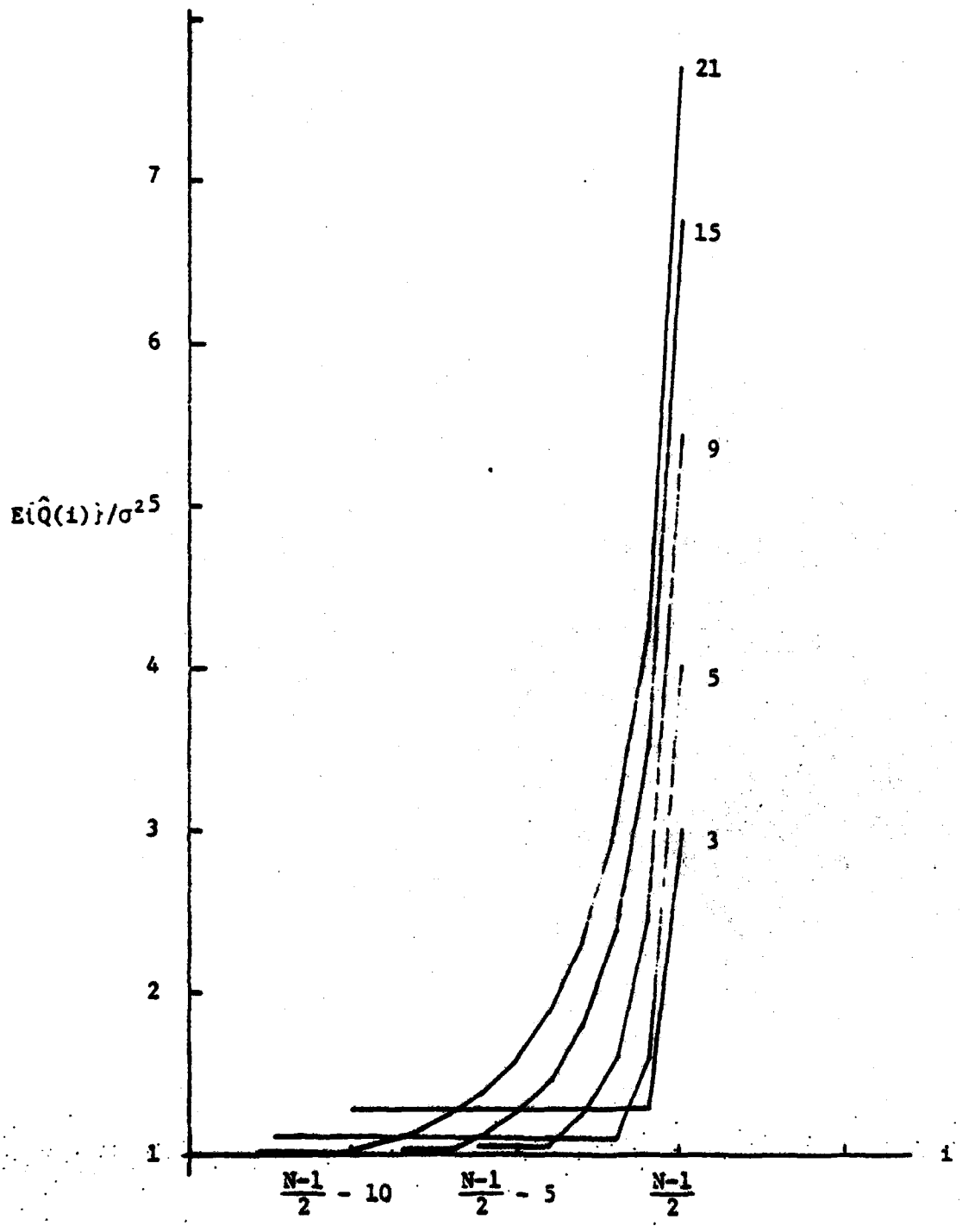


Figure B3. Approximate normalized robust variance estimate for various window sizes, S/σ large, using data from [20]. The estimate is symmetric around $i = N/2$.

$$E\{\hat{Q}(i)\} \simeq \frac{N}{(N-i)(N-i-1)} \sum_{j}' (w(j) - \tilde{w}(i))^2 \qquad \text{B-14}$$

where there are N-i terms in the sum. Results in [20] can be used to show that this can be written

$$E\{\hat{Q}(i)\} = \frac{N(N-i-2)}{(N-i)(N-i-1)} [1 + \frac{N-i}{N-i-2} E\{\tilde{w}^2_{p,N-i}\}/\sigma^2]\sigma^2 \qquad \text{B-15}$$

where $E\{\tilde{w}^2\} = \sigma^2_{\tilde{w}} + E^2\{\tilde{w}\}$, and $\tilde{w}_{p,\ N-i}$ is the $p^{\underline{th}}$-order element of the N-i near-zero valued elements, and $p=(N+1)/2 - i$, $1 \le i \le (N-1)/2$. That is, p is the median order of the N samples, but the median is among the N-i samples and has lesser order among them.

The data in Figure B3 is a plot of values of the formula B-15 for various window sizes N. The plot for N=5 is square-rooted and shown in Figure 2(e). The larger the window, the greater the ratio S/σ required for accuracy or relevancy of these plots. The peak moment $E\{\tilde{w}^2_{p,N-i}\}$ occurs when $i = \frac{N \pm 1}{2}$, as it is the 2nd moment of the first order statistic in the N-i set. For N=21 this moment is $7.7\sigma^2$. As the corresponding median, shown in Figure B1, is about 1.6σ, the variance of the N-i samples around their median is evidently about $7.7-1.6^2$, or a standard deviation of 2.27σ. If we require 5 deviations around 1.6σ to be less than S/2, then $(1.6 + 5(2.27))\sigma < S/2$, or $S/\sigma > 26$. Such a requirement is of course only necessary for the approximation of Equation B-15 to hold.

We note that any such estimate using Equation (26) with pulse-like errors as in Figure B3 could easily be smoothed with a following median filter. It is well known that median filters of length N_o will remove pulses of duration $(N_o-1)/2$ or less. Of course, any true step in noise variance would be detected, but at the cost of a time delay equal to half the length of the following filter.

REFERENCES

1. Carter, G.C. (editor), IEEE Transactions on Acoustics, Speech and Signal Processing, Part II, Special Issue on Time Delay Estimation. Vol. ASSP-29, No. 3, June 1981.

2. Carter, G.C., "Passive Sonar Signal Processing," IEEE Transactions ASSP, Vol. ASSP-29, No. 3, pp. 463-470. June 1981.

3. Schultheiss, P.M., and E. Weinstein, "Lower Bounds on the Localization Errors of a Moving Source Observed by a Passive Array," IEEE Transactions ASSP, Vol. ASSP-29, No. 3, pp. 600-607. June 1981.

4. J.C. Hassab and R.E. Boucher, "An Experimental Comparison of Optimum and Sub-optimum Filters' Effectiveness in the Generalized Correlator," J. Sound and Vibration, Vol. 76, pp. 117-128. 1981.

5. Hassab, J.C., B.W. Guimond, and S.C. Nardone, "Estimation of Location and Motion Parameters of a Moving Source Observed from a Linear Array," J. Acoustic Soc. America, Vol. 70, No. 4, pp. 1054-1061. October 1981.

6. Bradley, J.N., and R.L. Kirlin, "Delay Estimation by Expected Value," IEEE Transactions ASSP, Vol. ASSP-32, No. 1, pp. February 1984.

7. Weiss, A.J., and E. Weinstein, "Fundamental Limitations in Passive Time Delay Estimation - Part I: Narrow-band Systems," IEEE Transactions ASSP, Vol. ASSP-31, No. 2, pp. 472-486. April 1983.

8. Godiwala, P.M., "Passive Estimates of Underwater Maneuvering Targets," M.S. Thesis, Electrical Engineering Department, Virginia Polytechnic Institute, Blacksburg, VA. 1983.

9. Groutage, F.D., "State Variable Estimation using Adaptive Kalman Filter with Robust Smoothing," Proceedings 22nd IEEE Conference on Decision and Control, pp. , December 1983. IEEE No. 0191-2216/83/0000-0304.

10. Groutage, F.D., "Adaptive Robust Sequential Estimation with Application to Tracking a Maneuvering Target," Ph.D. Dissertation, Electrical Engineering Department, University of Wyoming, Laramie, WY 1982.

11. Myers, K.A. and B.D. Tapley, "Adaptive Sequential Estimation with Unknown Noise Statistics," IEEE Transactions Automatic Control, Vol. AC-21, No. 4, pp. 520-523. August 1976.

12. Boncelet, C.G. Jr., and B.W. Dickinson, "An Approach to Robust Kalman Filtering," Proceedings 22nd IEEE Conference on Decision and Control, pp. 304-305. December 1983. IEEE No. 0181-2216/83/0000- .

13. Launer, R.L. and G.N. Wilkinson (editors), Robustness in Statistics, Academic Press, New York. 1979.

14. Masreliez, C.J. and R.D. Martin, "Robust Bayesian Estimation for the Linear model and Robustifying the Kalman Filter," IEEE Trans. Automatic Control, Vol. AC-22, No. 3, June 1977, pp. 361-371.

15. Meyr, H., and G. Spies, "The Structure and Performance of Estimators for Real-Time Estimation of Randomly Varying Time Delay," IEEE Trans. ASSP, Vol. ASSP-32, No. 1, pp. 81-94.

16. Huang, T.S. (editor), Two Dimensional Digital Signal Processing II, Transforms and Median Filters, Springer-Verlag, N.Y., 1981.

17. Gallagher, N.C. Jr., and G.L. Wise, "A Theoretical Analysis of the Properties of Median Filters," IEEE Transactions ASSP, Vol. ASSP-29, No. 6, pp. 1136-1141. December 1981.

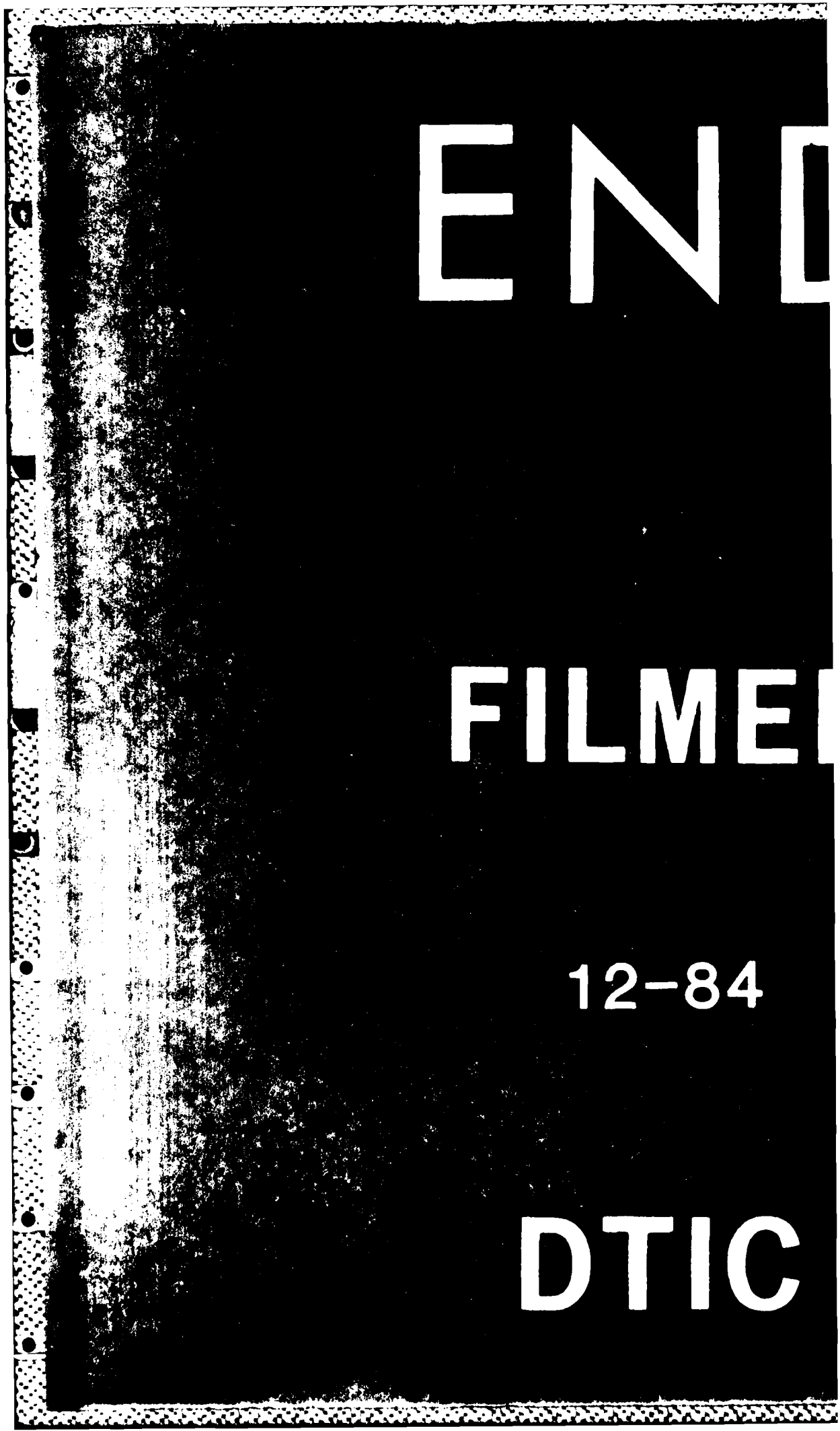
18. Mosteller, F. and J.W. Tukey, Data Analysis and Regression, Addison-Wesley, Reading, MA. 1977.

19. Gelb, A. (editor), Applied Optimal Estimation, MIT Press, Cambridge, MA. 1974.

20. Bovik, A.C., T.S. Huang and D.C. Munson Jr., "A Generalization of Median Filtering Using Linear Combinations of Order Statistics," IEEE Transactions ASSP, Vol. ASSP-31, No. 6, pp. 1342-1350. December 1983.

21. David, H.A., Order Statistics, Wiley, New York. 1980.

22. Teichroew, D. (editor), "Tables of Expected Values of Order Statistics and Products of Order Statistics for Samples of Size Twenty and Less from the Normal Distribution," Annals of Mathematical Statistics, Vol. 27, pp. 410-426. 1956.


END
FILME
12-84
DTIC